



# Forever Young: Aging Control For Smartphones in Hybrid Networks

Eitan Altman, Rachid El-Azouzi, Daniel Sadoc Menasche, Yuedong Xu

## ► To cite this version:

Eitan Altman, Rachid El-Azouzi, Daniel Sadoc Menasche, Yuedong Xu. Forever Young: Aging Control For Smartphones in Hybrid Networks. MobiHoc 2019 - 20th International Symposium on Mobile Ad Hoc Networking and Computing, Jul 2019, Catania, Italy. pp.91-100, 10.1145/3323679.3326507 . hal-02297281

**HAL Id: hal-02297281**

**<https://inria.hal.science/hal-02297281>**

Submitted on 9 Dec 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Forever Young: Aging Control For Smartphones in Hybrid Networks

Eitan Altman

INRIA & University of Avignon, France

Rachid El-Azouzi

University of Avignon, France

Daniel S. Menasché

UMass, USA; UFRJ, Brazil

Yuedong Xu

INRIA & University of Avignon, France

## Abstract

The demand for Internet services that require frequent updates through small messages, such as microblogging, has tremendously grown in the past few years. Although the use of such applications by domestic users is usually free, their access from mobile devices is subject to fees and consumes energy from limited batteries. If a user activates his mobile device and is in range of a service provider, a content update is received at the expense of monetary and energy costs. Thus, users face a tradeoff between such costs and their messages aging. The goal of this paper is to show how to cope with such a tradeoff, by devising *aging control policies*. An aging control policy consists of deciding, based on the current utility of the last message received, whether to activate the mobile device, and if so, which technology to use (WiFi or 3G). We present a model that yields the optimal aging control policy. Our model is based on a Markov Decision Process in which states correspond to message ages. Using our model, we show the existence of an optimal strategy in the class of threshold strategies, wherein users activate their mobile devices if the age of their messages surpasses a given threshold and remain inactive otherwise. We then consider strategic content providers (publishers) that offer *bonus packages* to users, so as to incent them to download updates of advertisement campaigns. We provide simple algorithms for publishers to determine optimal bonus levels, leveraging the fact that users adopt their optimal aging control strategies. The accuracy of our model is validated against traces from the UMass DieselNet bus network.

## 1 Introduction

The demand for Internet services that require frequent updates through small messages has tremendously grown in the past few years. While the popularity of traditional applications of that kind, such as weather forecasts, traffic reports and news, is unlikely to decline, novel applications, such as Twitter [1], have arisen. Twitter, alone, recorded a 1,500% increase in the number of registered users since 2006, and currently counts with more than 100 million users worldwide. For a second example, mobile messaging with Exchange ActiveSync [2] allows smartphone users to receive timely updates when new data items arrive in their mailboxes. Today, Exchange ActiveSync is supported by more than 300 million mobile devices worldwide. Henceforth,

for the sake of concreteness we focus on microblogging applications, such as Twitter or similar news feeds.

Users of microblogging applications join interest groups and aim at receiving small messages from editors. As messages age, they get outdated and their utilities decrease. As a consequence, users must control when to receive updates. A user willing to receive an update activates his mobile device, which then broadcasts periodic beacons to inform demands to service providers.

Although the use of microblogging applications by domestic users is usually free, their access from mobile devices consumes energy from limited batteries and is subject to fees. We consider users that can access the Internet either through a WiFi or 3G network. The 3G network provides broader coverage, but its usage requires a subscription to a cell phone data plan. More generally, our results apply to any scenario in which users can access the network through multiple interfaces with different costs and ubiquitousness [3, 4].

Let the *age* of a message held by a user be the duration of the interval of time since the message was downloaded by such user. If a user activates his mobile device and is in the range of a service provider (WiFi access point or 3G antenna), an update is received and the age of the message held by the user is reset to one, at the expense of the previously mentioned monetary and energy costs. Thus, users face a tradeoff between energy and monetary costs and their messages aging. To cope with such a tradeoff, users decide, based on the age of the stored message, whether to activate the mobile device, and if so, which technology to use (WiFi or 3G). We refer to a policy which determines activation decisions as a function of message ages as an *aging control policy*. *The first goal of this paper is to devise efficient aging control policies.*

Strategic content providers can incentivize users to download updates of advertisement campaigns or unpopular content by offering *bonus packages*. The goal of the bonus package, translated in terms of our aging control problem, consists of minimizing the average age of content held by users, subject to a budget on the number of messages transmitted per time slot, as dictated by the service provider capacity. Although nowadays bonus packages are set exclusively by service providers [5], we envision that in the future content providers will reach agreements with service providers. Through such agreements, content providers, also known as publishers, will play an important role in the settlement of bonus packages [6]. *The second goal of this paper is to solve the publishers' bonus selection problem, when users adopt an aging control policy.*

We pose the two following questions,

1. what is the users optimal aging control policy?
2. leveraging the users optimal aging control policy, what is the publishers optimal bonus strategy?

We propose a model that allows us to answer the questions above. Our model accounts for energy costs, prices and the utility of messages as a function of their age. Using our model, we show that users can maximize their utilities by adopting a simple threshold policy. The policy consists of activating the mobile device if the content age surpasses a given threshold and remaining inactive otherwise. We derive properties of the optimal threshold, and a closed-form expression for the average reward obtained by users as a function of the selected strategies. We then show the accuracy of our approach using traces collected from the UMass DieselNet

bus network. Using traces, we also study location-aware policies, according to which users can activate their devices based on their position on campus, and compare them against location-oblivious policies.

For the strategic publishers, we present two simple algorithms to solve the bonus determination problem posed above. The first algorithm presumes complete information while the second consists of a learning algorithm for publishers that have imperfect information about the system parameters, and is validated using trace-driven simulations. Finally, we show the convergence of the proposed learning algorithm, making use of results on differential inclusions and stochastic approximations.

In summary, we make the following contributions.

**Model formulation:** We introduce the aging control problem, and propose a model to solve it. Using the model, we derive properties about the optimal aging control policy and closed-form expressions for the expected average reward.

**DieselNet trace analysis:** We quantify how aging control policies impact users of the DieselNet bus network. Using traces collected from DieselNet, we show the accuracy of our model estimates and analyze policies that are out of the scope of our model.

**Mechanism design:** We provide two simple algorithms for publishers to incent users to download advertisement updates, leveraging the fact that users adopt their optimal aging control strategies. We formally show the convergence of the proposed learning algorithm, and numerically investigate its accuracy and convergence rate using trace-driven simulations.

The remainder of this paper is organized as follows. After further discussing the need for aging control, in §3 we present our model, and in §4 we report results obtained with traces from the UMass DieselNet bus network. Our model analysis is shown in §5, followed by the solution of the publishers problem in §6. We present a discussion of our modeling assumptions in §7, related work in §8 and §9 concludes the paper.

## 2 Why Aging Control?

The goal of an aging control policy is to provide high quality of service while 1) reducing energy consumption and 2) reducing 3G usage, by leveraging WiFi connectivity when available. Whereas the reduction in energy consumption is of interest mainly to subscribers, the reduction in 3G usage is of interest both to service providers and subscribers. Next, we discuss a few issues related to the adoption of aging control from the service providers and subscribers standpoints.

### 2.1 Service Provider Standpoint: Limited Spectrum

The increasing demand for mobile Internet access is creating pressure on the service providers, whose limited spectrum might not be sufficient to cope with the demand [7, 3]. To deal with such pressure, some wireless providers are offering incentives to subscribers to reduce their 3G usage by switching to WiFi [8]. Therefore, it is to the best interest of the service providers to devise efficient aging control policies for their users. Aging control policies can not only reduce the pressure on the 3G spectrum, but also reduce the costs to service

	Coverage	Energy Consumption (Scanning and Handshaking)	Monetary Costs
WiFi	10%-20%	0.63J (Android G1) 1.18J (Nokia N95)	free
3G	80%-90%	$\approx 0$	0.20 USD (SMS in USA)

Table 1: Coverage, energy consumption and monetary costs

providers that support hybrid 3G/WiFi networks (with savings of up to 60% against providers that offer only 3G [3]).

## 2.2 Users Standpoint: Energy Consumption, Monetary Costs and Coverage

How do the coverage, energy consumption and monetary costs vary between 3G and WiFi? Table 1 illustrates the answer with data obtained from [3, 9, 10, 11] and is further discussed next.

**Energy consumption and AP scanning** The energy efficiency of WiFi and 3G radios differs significantly. WiFi users actively scan for APs by broadcasting probes and waiting for beacons sent by the APs. If multiple beacons are received, users connect to the AP with highest signal strength. The association overhead, incurred when searching and connecting to the APs, yields substantial energy costs. Niranjana *et al.* [9] report that the energy consumption of a WiFi AP scan is 0.63J in Android G1 and 1.18J in Nokia N95, taking roughly 1 and 2 seconds, respectively (note that in one time slot users might perform multiple AP scans). In the remainder of this work, to simplify presentation, we consider only the energy costs related to association overhead. In many scenarios of practical interest WiFi and 3G radios incur comparable energy costs after the handshaking phase, and our model can be easily adapted to account for cases in which such energy costs differ.

**Monetary costs** Whereas free WiFi hotspots are gaining popularity [12, 13], the use of 3G is still associated to monetary costs. For example, in the United States and in Australia, it typically costs 0.20 USD to send an SMS through a 3G network [10]. As a second example, in Brazil subscribers of Claro 3G incur a cost of USD 0.10 per megabyte after exceeding their monthly quota [14].

**Coverage** It is well known that the coverage of 3G is much broader than the coverage of WiFi. This is because the 3G towers are placed by operators so as to achieve almost perfect coverage, while WiFi access points are distributed and activated in an ad hoc fashion. For instance, it has been reported in [3] that 3G and WiFi are available roughly 12% and 90% of the time in the town of Amherst, Massachusetts.

**Approximations** In light of the above observations, in the rest of this paper, except otherwise stated, we consider the following three approximations concerning energy consumption, monetary costs, and coverage. We assume that 1) the energy consumption of scanning for WiFi access points dominates the energy costs of the 3G and WiFi radios, 2) WiFi access points are open and freely available, whereas the use of 3G incurs a monetary cost per message and 3) WiFi is intermittently available whereas 3G offers perfect coverage. Note that the above three approximations are made solely to simplify presentation: our model has flexibility to

account for different energy consumption and monetary costs of WiFi and 3G, and can be easily adapted to account for a 3G network that does not have perfect coverage.

## 2.3 Adopting Aging Control

Next, we discuss two key aspects pertaining the adoption of aging control: 1) the delay tolerance of the applications and 2) the availability of WiFi access points.

### 2.3.1 Delay Tolerant Applications: Age and Utility

Aging control is useful for applications that can tolerate delays, such as news feeds and email. For these applications, users might be willing to tolerate some delay if that translates into reduced energy consumption or monetary costs.

**Workload** The workload subject to aging control can be due to one or multiple applications.

**One application:** Websites such as Yahoo! provide news feeds on different topics, ranging from business and economics to entertainment. Niranjana *et al.* [9] monitored feeds in ten categories, and reported that in three of these categories (business, top stories and opinion/editorial) at least one new feed is available every minute, with very high probability [9, Figure 13].

**Multiple applications:** Aiming at energy savings, it is common practice for users of smartphones to synchronize the updates of multiple applications, using APIs such as OVI Notifications [15]. In this case, the larger the number of applications that require updates, the higher the chances of at least one update being available every time slot.

In the remainder of this paper, we focus on the above mentioned workloads, for which updates are issued with high frequency. The analysis of light and medium workloads is subject for future work (see §7).

**User Interface** How does the utility of a content degrade with its age? The answer to this question is user and application dependent. It turns out that some applications, such as JuiceDefender [16], already allow iPhone users to specify their delay tolerance per application so as to save battery. For instance, users can simply specify a delay threshold, after which they wish to have received an update. In our model, this corresponds to an utility that has the shape of a step function (see §5.3).

### 2.3.2 WiFi Access Points

Open WiFi access points are widespread [12, 13]. The age control mechanisms described in this paper can rely on such access points, which users encounter in an ad hoc fashion. Alternatively, Internet service providers can deploy their own closed networks of WiFi access points, in order to alleviate the load of their 3G networks. In France, enterprises such as SFR and Neuf are already adopting such strategy. The age control mechanisms presented in this paper can be applied in this setting as well. In this case, users are required to perform a web-based login in order to access the WiFi spots, which can be done automatically using tools such as

Freewifi [17]. Note that even if mobile users have perfect geographic information about the location of the access points, random factors like attenuation (fading) and user speed will determine the availability of the access points, which will vary temporally and spatially.

**Markov Decision Processes** Remaining inactive at a given instant of time is particularly useful if future transmission opportunities are available in upcoming time slots. However, as discussed above, WiFi access points are intermittently available. Therefore, WiFi users experience a random delay between the instant of time at which they activate their devices and the instant at which they obtain content updates. How to account for the unpredictability of transmission opportunities in upcoming slots? To answer this question, in the next section we propose a model, based on a Markov Decision Process, which naturally accounts for the effect of the actions at a given time slot on the future states of the system, and allows us to derive the structure of the optimal aging control policy.

### 3 Model

We consider mobile users that subscribe to receive content updates. In the microblogging jargon, such users are said *to follow* a content. Content is transmitted from publishers to users, through messages sent by the service providers. The *age* of the message held by a user is defined as the length of time, measured in time slots, since the message was downloaded. Note that we assume that updates are available at every time slot with high probability (see §7).

Let  $x_t$  be the age of the message held by a user at time  $t$ . The age of the message equals one when the user first receives it, and increases by one every time slot, except when the user obtains an update, time at which the age is reset to one.

A user can receive message updates when his mobile device is in range of a service provider and the contact between them lasts for a minimum amount time which characterizes a *useful contact opportunity*. While the 3G technology is assumed to guarantee perfect coverage, WiFi users are subject to outages. Let  $e_t$  be an indicator random variable equal to 1 if there is a useful contact opportunity with a WiFi provider at time  $t$ , and 0 otherwise. We let  $p = E[e_t]$  and assume  $0 < p < 1$ . Next, we state our key modeling assumption.

**Assumption 3.1. *Uniform and independent contact opportunity distribution:*** *The probability of a useful contact opportunity between a user and WiFi providers is constant and independent across time slots, and equals  $p$ .*

Under Assumption 3.1, there are no correlations in time between contact opportunities experienced by a user; this is a strong assumption, since such correlations are present in any mobile network, as illustrated in §4.1. However, as shown in §4.2, there are scenarios of practical interest in which the uniformity and independence assumption does not compromise the accuracy of the results obtained using our model. There-

fore, we proceed our analysis under such an assumption, indicating its implications throughout the paper and studying scenarios that are out of the scope of our model using real-world traces.

The goal of each user is to minimize the expected age of the content he follows, accounting for energy and monetary costs. In order to achieve such goal, users must choose at each time slot  $t$  their actions,  $a_t$ . The available actions are

$$a_t = \begin{cases} 0, & \text{inactive} \\ 1, & \text{WiFi} \\ 2, & \text{WiFi if } \exists \text{ useful contact opportunity, else 3G} \end{cases}$$

**State Dynamics** The age of the content followed by a user increases by one if his device is inactive or if it is not in range of a service provider, and is reset to one otherwise. Let  $M$  be the maximum age of a content. Then,

$$x_{t+1} = \begin{cases} \min(x_t + 1, M), & \text{if } (a_t=0) \text{ or } (a_t=1 \text{ and } e_t=0) \\ 1, & \text{if } (a_t=2) \text{ or } (a_t=1 \text{ and } e_t=1) \end{cases} \quad (1)$$

**Utility** Let  $U(x_t)$  the utility of the followed content at time  $t$ . We assume that  $U(x)$  is a non-increasing function of  $x$ , which corresponds to messages that become obsolete with time, and that  $U(x) = Z$  if  $x \geq M$ .

**Costs** Let  $G$  be the cost incurred to maintain the mobile device active, measured in monetary units. Then, the energy cost  $c_t$  is given as a function of  $a_t$  as

$$c_t(a_t) = \begin{cases} G, & \text{if } a_t \geq 1 \\ 0, & \text{if } a_t = 0 \end{cases} \quad (2)$$

Service providers charge a price for each message transmitted. The prices charged by WiFi and 3G providers are  $P$  and  $P_{3G}$ , respectively. When a user receives an update, he is subject to a monetary cost of  $m_t$ ,

$$m_t(a_t, e_t) = \begin{cases} P_{3G}, & \text{if } a_t = 2 \text{ and } e_t = 0 \\ P, & \text{if } a_t \geq 1 \text{ and } e_t = 1 \end{cases} \quad (3)$$

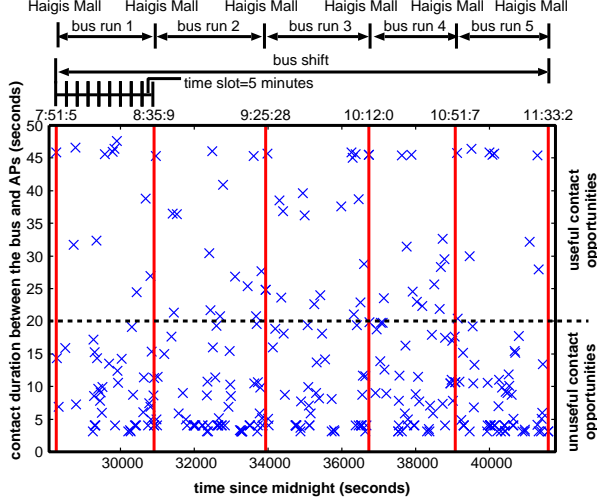
Content providers, also referred to as publishers, can offer bonuses to users that follow their contents. Such bonuses are set in agreement with service providers, and are transfered to users as credits. Let  $B_t$  be the bonus level set by the content provider,  $B_t \leq \min(P_{3G}, P)$ .

The instantaneous user reward at time  $t$ ,  $r_t(x_t, a_t)$ , is

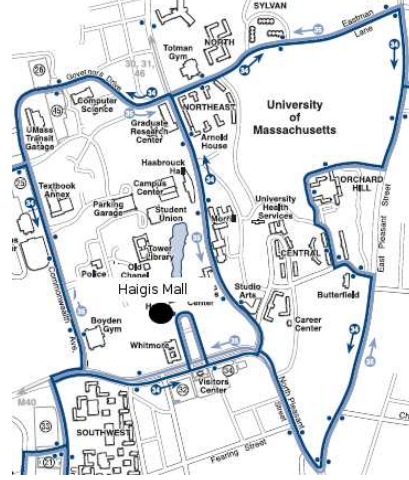
$$r_t(x_t, a_t) = U(x_t) - c_t(a_t) - \max(m_t - B_t, 0) \quad (4)$$

**Users Strategies** The *strategy* of a user is given by the probability of choosing a given action  $a_t$  at each of the possible states. Without loss of generality, in this paper we restrict to Markovian stationary policies [18,





(a) Bus-AP contacts during a typical bus shift



(b) Route of a bus that passes through Haigis Mall

Figure 1: A typical bus shift.

Chapter 8]. The probability of choosing action  $a_t$  at state  $x_t$  is denoted by  $u(a_t|x_t)$ .

**Problem Definition** The problem faced by each user consists of finding the strategy  $u$  that maximizes his expected average reward.

USER PROBLEM: Obtain strategy  $u$  so as to maximize  $E[r; u]$ , where

$$E[r; u] = \lim_{\ell \rightarrow \infty} \frac{1}{\ell} \sum_{t=0}^{\ell} E[r_t(x_t, a_t); u] \quad (5)$$

In what follows, we drop the subscript  $t$  from variables when analyzing the system in steady state.

**Optimal Threshold Policy** Next, we introduce two special classes of policies, the *two-threshold policies* and the *threshold policies*. Let  $s$  and  $s_{3G}$  be the WiFi and the 3G thresholds,  $1 \leq s \leq s_{3G} \leq M+1$ . A policy which consists of setting  $a = 0$  when  $x < s$ ,  $a = 1$  when  $s \leq x < s_{3G}$ , and  $a = 2$  if  $x \geq s_{3G}$ , is referred to as a *two-threshold policy*. If users have no access to 3G ( $P_{3G} \rightarrow \infty$ ), a policy which consists of backing off if  $x < s$  and being active if  $x \geq s$  is referred to as a *threshold policy*.

Note that while  $x \in [1, M]$ , the thresholds assume values in the range  $[1, M+1]$ . When  $P_{3G} \rightarrow \infty$ ,  $s = M+1$  means that the user should remain always inactive (refer to Table 2 for notation).

The following proposition reduces the problem of finding the optimal policy for the USER PROBLEM to the one of finding the two thresholds  $s$  and  $s_{3G}$ .

**Proposition 3.1.** *The USER PROBLEM admits an optimal policy of two-threshold type such that*

$$\begin{array}{c|c}
\begin{array}{l} \text{if } P_{3G} > G/p + P, \\ \\ a(x) = \begin{cases} 0, & \text{if } x < s \\ 1, & \text{if } s \leq x < s_{3G} \\ 2, & \text{otherwise} \end{cases} \end{array} & \begin{array}{l} \text{if } P_{3G} \leq G/p + P, \\ \\ a(x) = \begin{cases} 0, & \text{if } x < s_{3G} \\ 2, & \text{otherwise} \end{cases} \end{array}
\end{array}$$

If  $P_{3G} \rightarrow \infty$ ,  $s_{3G} = M + 1$  and the USER PROBLEM admits an optimal policy of threshold type.

When the price charged by the 3G providers dominates the costs, the optimal policy consists of activating the WiFi radio if  $s \leq x < s_{3G}$  and using the 3G only if  $x \geq s_{3G}$ . If the price charge by the 3G providers is dominated by  $G/p + P$ , though, users are better off relying on the 3G technology, which offers perfect coverage.

To simplify presentation, in the upcoming sections we assume that 1) users have access only to WiFi APs ( $P_{3G} \rightarrow \infty$ ) and 2) the optimal threshold policy is unique. In Appendix C we 1) show how our results extend to the scenario in which users can choose between WiFi and 3G and 2) characterize the scenarios under which the optimal threshold is not unique.

**Definition 3.1.** *The threshold  $s^*$  of the optimal policy is*

$$s^* = \operatorname{argmax}_s \{E[r; s]\} \quad (6)$$

The properties of the optimal threshold are discussed in the following section in light of traces from DieselNet, and are formally stated in §5.

## 4 Evaluating Aging Control Policies in DieselNet

In this section we use traces collected from the UMass Amherst DieselNet [19] to evaluate how aging control policies perform in practice. Our goals are 1) to show the accuracy of our model predictions, 2) to assess the optimality of the model predictions in the class of threshold policies and 3) to compare the optimal policy obtained with our model against policies that are out of the scope of our model, such as location-aware policies.

The users of the microblogging application considered in this section are passengers and drivers of buses. We assume that users are interested in following one of the heavy load news feeds described in §2.3.1, for which updates are issued every minute with high probability. Time is divided into slots of 5 minutes.

We begin with an overview of statistics collected from the traces and their implications on aging control.

### 4.1 Measuring Contact Statistics

To characterize the update opportunities experienced by the users, we analyze contacts between buses and access points (APs) at the UMass campus. The traces were compiled during Fall 2007 from buses running

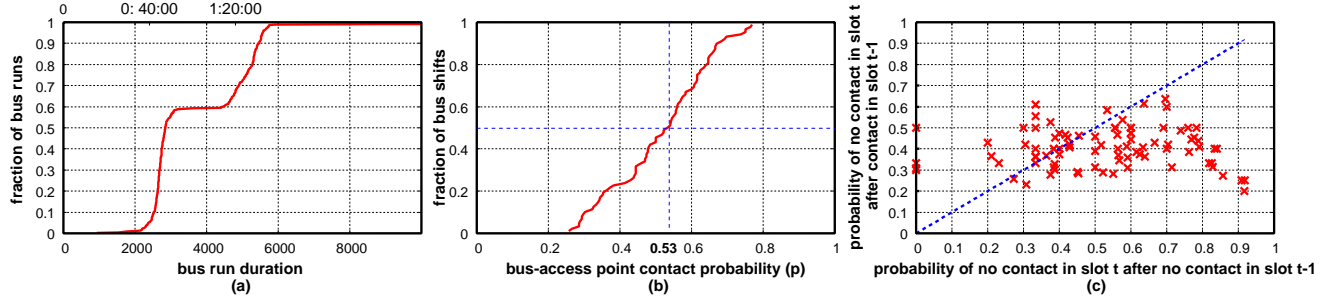


Figure 2: Useful contact opportunity statistics (a) time between contacts to AP at Haigis Mall; (b) CDF of  $p$ ; (c) scatter plot of opportunities in consecutive slots.

routes serviced by UMass Transit. Each bus scans for connection with APs on the road, and when found, connects to the AP and records the duration of the connection [19].

Figure 1(a) illustrates a sample of the trace data. In Figure 1(a), each cross corresponds to a contact between the bus and an AP. The  $x$  and  $y$  coordinates of each cross correspond to the time of the day at which the contact occurred and the duration of the contact, respectively.

In the rest of this paper, we assume that when the mobile devices are active, scans for access points occur every 20 seconds, and last up to the moment at which an access point is found. Ra *et al.* [11, §5] empirically determined that a scanning frequency of 1/20 seconds yields a good balance between efficiency and low energy expenditure. An access point is considered useful once it is scanned in two consecutive intervals of 20 seconds. Henceforth, we refer to contact opportunities that last at least 20 seconds as *useful contact opportunities*, or *contacts* for short, when the qualification is clear from the context. Time slots in which at least one useful contact opportunity begins are referred to as *useful slots*.

Figure 1(b) shows the map of one of the bus routes considered in this paper, with the Haigis Mall in evidence. The Haigis Mall is a central location for the transit of the buses at UMass, and in this work we restrict to routes that pass through it. In Figure 1(a), vertical lines correspond to instants at which the bus passed through the Haigis Mall. A *bus run* corresponds to the interval between two arrivals at Haigis Mall. A *bus shift* corresponds to a sequence of consecutive and uninterrupted bus runs, in the same day. Note that a bus run in Figure 1(a) takes around 40 minutes. A typical bus run varies between 40 minutes and 1 hour and 20 minutes (see Figure 2(a)).

Figure 1(a) shows that roughly every time the bus passes through Haigis Mall there is a useful contact opportunity. This observation has important implications on the aging control policy. In particular, it indicates that users that are location-aware can take advantage of such information in order to devise their activation strategies. We will evaluate the performance of the policy which consists of activating the mobile device only at Haigis Mall in §4.2.4.

Next, our goal is to study the distribution of contact opportunities between buses and APs. In particular, we wish to identify the extent at which the uniformity and independence assumption (Assumption 3.1) holds in practice, specially when buses are far away from Haigis Mall (see Figure 1). To this aim, for each day

and for each bus shift, we generate out of our traces a *string of ones and zeros, corresponding to useful and unuseful slots*, respectively. Such strings are used to plot Figures 2(b) and 2(c), as well as an input to our trace-driven simulator, described in the next section.

Figure 2(b) shows the CDF of the bus-AP contact probability ( $p$ ) across bus shifts. The median of  $p$  is 0.53. The contact probability is above 0.3 for up to 85% of the bus shifts and below 0.7 for more than 90% of the bus shifts.

Figure 2(c) shows a scatter plot of the contact probabilities in two consecutive slots. Each cross corresponds to a bus shift. A cross with coordinates  $x$  and  $y$  corresponds to a bus shift in which the probability of no contact in slot  $t$  after no contact in slot  $t - 1$  is  $x$  and the probability of no contact in slot  $t$  after a contact in slot  $t - 1$  is  $y$ . The figure indicates that when  $x$  varies between 0.3 and 0.7, a significant fraction of points is close to the line  $x = y$ . This behavior is similar to the one expected in case contacts are approximately uniform and independent from each other.

## 4.2 Evaluating Aging Control Policies

In this section we validate our model against traces. We begin by describing our methodology and reference configuration, and then consider both location-oblivious and location-aware policies.

### 4.2.1 Methodology and Reference Configuration

To validate our model we use the traces described in the previous section. We assume that users do not have a cell phone data plan ( $P_{3G} \rightarrow \infty$ ) and WiFi is free ( $P = B = 0$ ).

The computation of the optimal policy using the proposed model requires estimates of  $p$ ,  $U(x)$  and  $G$ . For a given bus shift,  $p$  is estimated as the number of useful slots in that shift divided by the total number of slots. Note that to compute the optimal strategy using our model we assume knowledge of  $p$ , but not of the distribution of contact times and durations. When searching for the optimal threshold strategy using traces, in contrast, we perform trace-driven simulations. Our simulator, as well as extensive statistics obtained from the traces, are available at <http://www-net.cs.umass.edu/~sadoc/agecontrol/>. For each bus shift, our simulator takes as input the string of ones and zeros corresponding to useful and unuseful slots, and computes the reward experienced by a user adopting a given activation policy. The strategy that yields the highest reward correspond to the *optimal trace-driven policy*.

In our reference setting, the utility of messages decays linearly during a bus run, and remains zero afterwards,  $U(x) = \max(M - x, 0)$ . We vary  $M$  between 10 and 16 (which correspond to 50 and 80 minutes, resp., see also Figure 2(a)), in increments of 2. For ease of presentation, let  $b$  be the energy cost scaled by a factor of  $1/(M - 1)$ ,

$$b = G/(M - 1) \tag{7}$$

We vary  $b$  according to our experimental goals between 0.2, 0.8 and 1.8, corresponding to small, medium and high costs, respectively.

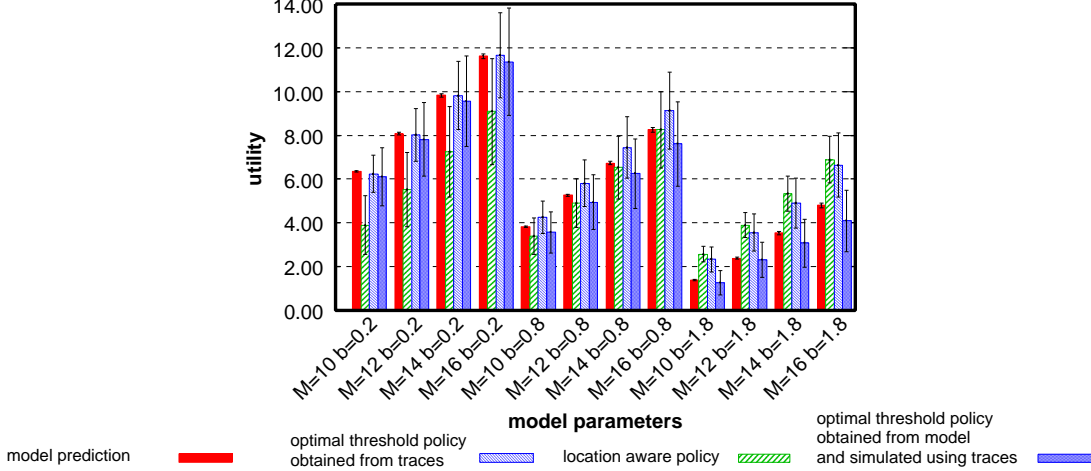


Figure 3: Rewards as predicted by our model and observed in DieselNet.

#### 4.2.2 Location Oblivious Policies

Next, we validate our model against traces assuming that users are restricted to threshold policies (non-threshold policies are considered in the next section). Figure 3 shows, for different model parameters, the averages of 1) the optimal reward *predicted* by our model, 2) the reward *effectively obtained* using the optimal strategy computed by the model as input to our trace-driven simulator and 3) the *optimal trace-driven (threshold) policy*, with the 95% confidence intervals (to increase the number of samples, each bus shift is replayed 40 times).

In Figure 3, note that across all parameters, the rewards predicted by our model match the rewards effectively obtained using the policy proposed by our model pretty well. When the energy cost is low ( $b = 0.2$ ), our model predictions also closely match the optimal trace-driven policy. When the energy cost is medium ( $b = 0.8$ ) the accuracy of the predictions of our model depends on the maximum age  $M$ . Recall that typical bus runs last between 40 minutes and 1 hour and 20 minutes (see Figure 2(a)) and that when a bus run is completed, a contact occurs with high probability (see Figure 1(a)). If the maximum age is 50 minutes ( $M = 10$ ), the fact that our model does not capture the correlations among contacts between buses and APs does not play an important role. However, as  $M$  increases, strategically setting the policy to account for such correlations is relevant. Similar reasoning holds when  $b = 1.8$ .

Figure 4(a) shows the distribution of the distance between the optimal threshold using our model versus the optimal trace-driven policy, for different model parameters. In accordance to our previous observations, when  $b = 0.2$ , the distance between the optimal threshold and the one computed by our model is smaller than or equal to two in at least 60% of the bus shifts. The same holds when  $b = 0.8$  and  $M = 10$ . When  $M = 16$  and  $b = 1.8$ , in contrast, the distance to the optimal threshold is smaller than two for less than 40% of the bus shifts.

Figure 4(b) reports the mode of the optimal threshold as estimated by our trace-driven simulator and by our model. The mode of the optimal threshold increases with the energy cost (see Proposition 5.3), and the

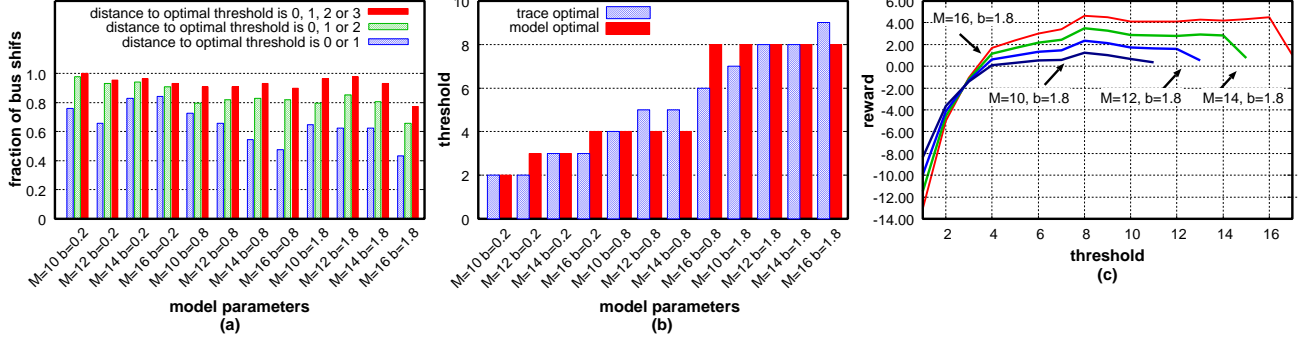


Figure 4: Model validation (a) distance between optimal trace-driven threshold and optimal model threshold; (b) optimal threshold mode; (c) trace-driven reward as a function of the threshold, where threshold is assumed the same at all bus shifts.

distance between the simulator and model predictions never surpasses two.

Figure 4(c) provides further insight on the problem when  $b = 1.8$ . Until this point, we allowed different bus shifts to correspond to different threshold policies (i.e., we considered bus shift discriminated strategies). Next, we consider users that adopt the same threshold policy over all bus shifts (i.e., we consider flat strategies over bus shifts). For different threshold values, Figure 4(c) shows the average reward obtained using our trace-driven simulations (see Figure 3 for confidence intervals). Note that the optimal threshold equals eight in the four scenarios under consideration. This threshold value, in turn, is in agreement with the mode predicted by our model (see Figure 4(b)). In Figure 5(a) we reproduce the results of Figure 4(c) using our model. We let  $p = 0.53$ , the median contact probability (see Figure 2(c)). Comparing Figure 4(c) and Figure 5(a), we observe that the empirical curves are predicted by our model with remarkable accuracy. The distance between the optimal threshold predicted by our model (marked with circles in Figure 5(a)) and the one obtained through the trace-driven simulations (consistently equal to eight) is smaller than or equal to 1, and the utility discrepancy does not surpass 0.5.

#### 4.2.3 Location Aware Policy

Users can leverage geographic information in order to decide when to activate their mobile devices. At UMass, for instance, a central bus stop is located at the Haigis Mall (see Figure 1(b)). Buses usually linger at Haigis Mall for a couple of minutes, time at which transmission opportunities usually arise (see Figure 1(a)). Therefore, in this section we consider the following location aware policy: activate the mobile device if in Haigis Mall, and remain inactive otherwise. Figure 3 shows the performance of our location aware policy obtained using trace-driven simulations. When the energy costs are not high ( $b = 0.2$  or  $0.8$ ), it is always advantageous to activate the mobile device before the bus returns to Haigis Mall, in order to opportunistically take advantage of contacts with APs during the bus run. However, when the energy costs are high ( $b = 1.8$ ), our location aware policy outperforms the best trace-driven threshold policy. The non-optimality of threshold policies in this scenario corroborates the fact that when energy costs are high, it is important to account for correlations among contact opportunities.

variable	description	
$a_t$	action of user	
$x_t$	age of message	
$r_t$	reward	
$e_t$	=1 in case of useful contact opportunity, =0 otherwise	
parameter	description	
$U(x)$	utility of message at age $x$	user
$B_t$	bonus level	publisher
$P_{3G}$	price of 3G	3G service provider
$P$	price of WiFi	WiFi service provider
$G$	activation cost	
$p$	probability of useful contact opportunity, $p = E[e_t]$	

Table 2: Table of notation. Notes: (1) Subscripts are dropped when in steady state. (2)  $B$  is a parameter in §5 and a variable in §6.

#### 4.2.4 Summary

To sum up, our model accurately predicts the reward effectively obtained by its proposed policy. If the energy costs are low, or if users do not discriminate their strategies with respect to the bus shifts, our model can also predict the optimal policy obtained through trace-driven simulations. Otherwise, users can benefit from correlations among contact opportunities.

## 5 Model Analysis

Our goals now are (a) to derive the optimality conditions that must be satisfied by the optimal policy; (b) to show properties of the optimal threshold and (c) to present specialized results for step and linear utility functions. We tackle each of the goals in one of the subsequent sections, respectively.

### 5.1 Optimal Policy General Structure

We now derive the general structure of the optimal policy. To this goal, consider a fixed policy  $u$ . In this section we assume that, under  $u$ , users have a positive activation probability in at least one state. The conditions for the optimality of the policy which consists of remaining always inactive will be established in Proposition 5.1.

Let  $\mathcal{P}_u$  denote the transition probability matrix of the Markov chain  $\{x_t : t = 1, 2, \dots\}$  which characterizes the dynamics of the age, given policy  $u$ . Let  $r_u(x)$  be the expected instantaneous reward received in a time slot when the system is in state  $x$  and policy  $u$  is used. The vector of expected instantaneous rewards is denoted by  $\mathbf{r}_u$ .

Let the gain,  $g_u$ , be the average reward per time slot in steady state,  $g_u = E[r; u]$ . Since the number of states in the system is finite and from each state there is a positive probability of returning to state 1,  $\mathcal{P}_u$  comprises a single connected component, and  $g_u$  does not depend on the initial system state.

The *relative reward* of state  $x$  at time  $t$ ,  $V(x, t)$ , is the difference between the expected total reward

accumulated when the system starts in  $x$  and the expected total reward accumulated when the system starts in steady state. Let  $V(x) = \lim_{\ell \rightarrow \infty} V(x, \ell) = \lim_{\ell \rightarrow \infty} \left( \sum_{t=0}^{\ell} \mathcal{P}_u^t(\mathbf{r}_u - g_u \mathbf{e}) \right)(x)$ , where  $\mathbf{e}$  is a column vector with all its elements equal to one. It follows from (4) and similar arguments as to those in [18, eq. (8.4.2)] that a policy which satisfies the following conditions, for  $1 \leq x \leq M$ , is optimal,

$$V(x) = \max \left( U(x) + V(\min(x+1, M)) - g_u, U(x) - G + p(V(1) - P + B) + (1-p)V(\min(x+1, M)) - g_u \right)$$

An optimal policy is obtained from the optimality conditions as follows,

$$a(x) = \begin{cases} 0, & \text{if } -G/p + (V(1) - P + B) \leq V(\min(x+1, M)) \\ 1, & \text{otherwise} \end{cases} \quad (8)$$

In Appendix A we show that  $V(x)$  is decreasing on  $x$  when  $P = 0$  (the corresponding result when  $P \geq 0$  can be found in Appendix A.3). Thus, the existence of an optimal policy in the class of threshold policies (Proposition 3.1) follows from (8).

Note that adding a constant to  $U(x)$ ,  $1 \leq x \leq M$ , does not affect the optimal policy (8). Therefore, in the rest of this paper we assume, without loss of generality, that  $U(M) = 0$ .

## 5.2 Optimal Threshold Properties

In this section we aim at finding properties of the optimal threshold. To this goal, we note that an user adopting a threshold strategy goes through cycles. Each cycle consists of an idle and active period. An idle period is initiated when the age is one, and ends immediately before the instant at which the age reaches the threshold  $s$ . At age  $s$ , an active period begins and lasts on average  $1/p$  time slots, up to the instant at which the age is reset to one.

The following proposition establishes conditions according to which the optimal actions are invariant in time (see Appendix Appendix A). The proof of the subsequent results are available in the appendices.

**Proposition 5.1.** *The optimal policy consists of being always active if and only if*

$$\frac{1}{1-p} \left( U(1) - p \sum_{x=1}^{M-1} U(x)(1-p)^{x-1} \right) \geq \frac{G}{p} + P - B \quad (9)$$

*The optimal policy consists of being always inactive if and only if*

$$\sum_{j=1}^{M-1} U(M-j) \leq \frac{G}{p} + P - B \quad (10)$$

Given that the condition for the optimal policy to be always inactive is established in the second part of Proposition 5.1, in what follows we focus on policies which consist of being active in at least one state.

For a fixed threshold policy with corresponding threshold  $s$ , let the system state transition probability matrix,  $\mathcal{P}_s$ , be



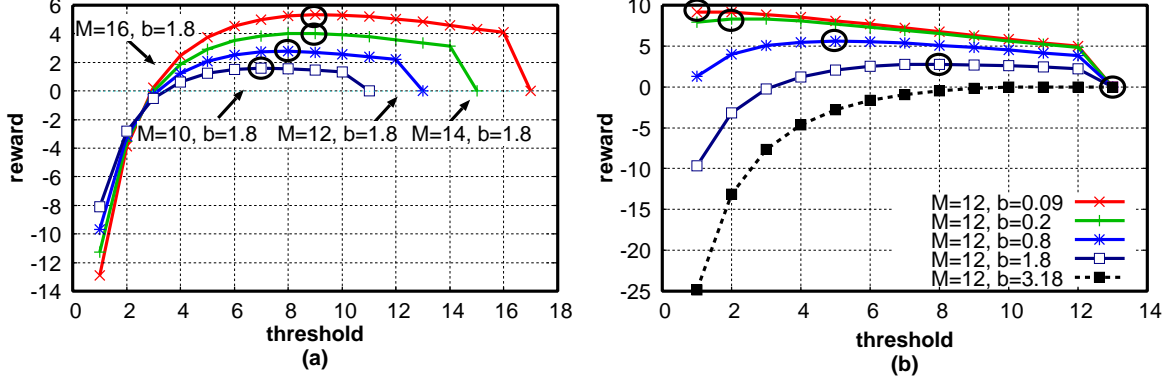


Figure 5: Model numerical evaluation ( $U(x) = M - x$ ,  $P = 0$ ,  $p = 0.54$ ).

$$\mathcal{P}_s = \begin{bmatrix} 1 & 0 & 1 & & & \\ & 0 & 0 & 1 & & \\ & & \ddots & & & \\ & & & 0 & 1-p & \\ & & & & \ddots & \\ p & & & & & 1-p \end{bmatrix} \quad (11)$$

Let  $\pi$  be the steady state solution of the system,  $\pi \mathcal{P}_s = \pi$ . Then, the fraction of time slots in which an user issues updates is  $\pi_1$  (see Appendix A.1),

$$\pi_1 = \left( s + \frac{1-p}{p} \right)^{-1} \quad (12)$$

Next, we derive a closed-form expression for the expected reward as a function of the threshold  $s$ . In Appendix A.2 we show that replacing the expression of the steady state solution  $\pi$  into  $E[r; s] = \sum_{i=1}^M \pi_i r(i, \mathbf{1}_{i \geq s})$  yields,

$$E[r; s] = \pi_1 \left[ \sum_{x=1}^{s-1} U(x) + \sum_{i=0}^{M-1-s} U(i+s)(1-p)^i - \frac{G}{p} - P + B \right]$$

The derivation of the equation above is found in Appendix A.2. In Appendix A.6 we use the expression above to show that the expected average reward is non-decreasing in  $s$ , for  $s < s^*$ , as stated in the following proposition.

**Proposition 5.2.** *The optimal threshold value  $s^*$  is*

$$s^* = \min \left\{ s \mid E[r; s] \geq E[r; s+1] \right\} \quad (13)$$

Finally, Proposition 5.3 formalizes the monotonicity of  $s^*$  with respect to  $G$  observed in §4.

**Proposition 5.3.** *The optimal threshold  $s^*$  is non-decreasing with respect to  $G$  and  $P$  and non-increasing with respect to  $B$ .*

### 5.3 Special Utility Functions

Next, we specialize our results to two classes of utilities.

#### 5.3.1 Step Utility

Let the utility function  $U(x)$  be given by a step function,  $U(x) = v$  if  $x \leq k$  and  $U(x) = 0$  otherwise. The next proposition allows us to efficiently compute the optimal threshold in this case. Let  $\varphi$  be the root of  $\frac{d}{ds}E[r; s] = 0$  when  $k \leq s$ . The closed form expression of  $\varphi$  is given in Appendix A.8.

**Proposition 5.4.** *For the step function utility, the optimal threshold  $s^*$  is equal to either 1,  $k - 1$ ,  $\lfloor \varphi \rfloor$ ,  $\lceil \varphi \rceil$ ,  $M$  or  $M + 1$ .*

#### 5.3.2 Linear Utility

Let  $U(x) = M - x$ . Then,  $E[r; s]$  can be expressed in closed form as a function of  $s$ ,  $M$ ,  $p$  and  $G$  (see Appendix A.9). To illustrate the behavior of  $E[r; s]$ , we let  $p = 0.54$  (see Figure 2),  $P = B = 0$ , and vary  $b$  and  $M$  as shown in the legend of Figure 5. Figure 5(a) was discussed in §4.2.2 in light of our trace-driven simulations. Figure 5(b) shows that when  $b = 0.09$  (resp.,  $b = 3.18$ ), inequality (9) (resp., (10)) holds and the optimal threshold is 1 (resp., 13), in agreement to Proposition 5.1. In accordance to Proposition 5.2, Figure 5(b) also indicates that when  $s < s^*$  the reward is increasing. Finally, the optimal threshold in Figure 5(b) increase as a function of  $b = G/(M - 1)$ , which serves to illustrate Proposition 5.3.

## 6 The Publisher Bonus Package

In this section we consider strategic publishers that offer bonus packages to users, so as to incent them to download updates of advertisement campaigns or unpopular content. In §6.1 we assume that publishers have complete information on the system parameters and consider the incomplete information case in §6.2.

### 6.1 Complete Information

Next, we consider publishers that, while devising their optimal bonus strategies, leverage the fact that users solve the USER PROBLEM. The optimal bonus strategy consists of finding the bonus level  $B$  that minimizes the average age of messages in the network, under the constraint that the expected number of messages transmitted per time slot is below a given budget, dictated by the service provider.

Let  $N$  be the number of users in the network. In what follows, we make the dependence of  $s^*$  and  $\pi_1$  (see (12)) on the bonus level  $B$  explicit. Let  $Q$  be the average number of messages transmitted per time slot.

$$Q = N\pi_1(B) = N/(s^*(B) + (1 - p)/p) \quad (14)$$

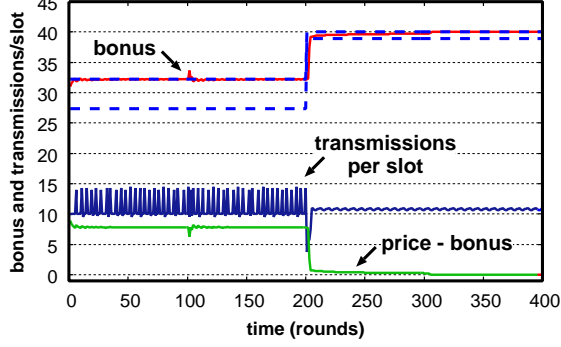


Figure 6: Learning algorithm: sample path of trace-driven simulation

Let  $A$  be the average age of messages in the network (see Appendix A.10 for its closed-form expression). Let  $T$  be the constraint imposed by the service provider on the expected number of messages sent per time slot. Then, the problem faced by the publisher is

PUBLISHER PROBLEM: Obtain bonus level  $B$  so as to

$$\min \quad A = \sum_{i=1}^M i \pi_i(B) \quad (15)$$

$$s.t. \quad Q \leq T \quad (16)$$

Note that since  $s^*(B)$  is not injective, there might be a range of bonuses levels that solves the PUBLISHER PROBLEM.

Under the assumption that the above problem admits a solution, which is guaranteed to exist if  $s(P) \geq N/T - (1-p)/p$ ,

**Proposition 6.1.** *The solution of the PUBLISHER PROBLEM consists of setting the bonus level  $B$  in such a way that  $s = \max(0, \min(M+1, \lceil N/T - (1-p)/p \rceil))$ . The solution can be found using a binary search algorithm.*

## 6.2 Incomplete Information: Online Learning Algorithm

How publishers can set their bonus levels without knowing the number of users in the system and their strategies? To answer this question, we present a simple learning algorithm to solve the PUBLISHER PROBLEM when the system parameters are unknown.

---

**Algorithm 1** Online estimation of optimal bonus level.

---

- 1: **Input:** maximum bonus level  $\widehat{B}$ , target number of messages per slot  $T$ , round duration  $\tau$  time slots, learning rate  $\alpha$
  - 2: Choose initial bonus level  $B_0$  such that  $B_0 \in [0, \widehat{B}]$ ;  $t \leftarrow 0$
  - 3: **while**  $|T - Q_t| > \epsilon$  **do**
  - 4:   At the end of round  $t$ ,
  - 5:    $Q_t \leftarrow R_t/\tau$
  - 6:    $B_{t+1} \leftarrow \min(\widehat{B}, \max(0, (B_t + \alpha(T - Q_t)/t)))$
  - 7:    $t \leftarrow t + 1$
  - 8: **end while**
-

The proposed algorithm proceeds in rounds. Each round corresponds to  $\tau$  time slots, at which users have their requests served. Let  $B_t$  be the bonus set by the publisher at the beginning of round  $t$ , and let  $R_t$  be the number of requests served at that round. The average number of requests served per time slot in round  $t$  is  $Q_t = R_t/\tau$ .

Algorithm 1 updates the bonus level as follows. At round  $t$ , if the average number of requests served per time slot is above the target  $T$ , the bonus level is decreased by  $\alpha(Q_t - T)/t$ ; otherwise, the bonus level is increased by  $\alpha(T - Q_t)/t$ . The learning step size,  $\alpha$ , is a learning parameter that impacts the algorithm convergence time. Smaller values of  $\alpha$  yield a smoother but slower convergence [20, Chapter 5]. The bonus level is required to be positive, and cannot surpass the maximum bonus level,  $\hat{B}$  (line 6 in Algorithm 1). The algorithm stops when  $|T - Q_t| \leq \epsilon$ , where  $\epsilon > 0$  is the tolerance parameter (line 3). Under the assumption that the system parameters (see Table 2) are fixed,

**Proposition 6.2.** *The sequence of bonus levels  $\{B_t\}_{t=0}^{\infty}$  converges to the optimal solution of the PUBLISHER PROBLEM with probability one.*

According to Proposition 6.2, the convergence of Algorithm 1 is assured when the system parameters are fixed. In order to 1) study the behavior of our algorithm when the population size varies and to 2) investigate how the convergence speed may be affected by the correlation among users, we conducted trace-driven simulations, whose results are reported in Appendix D.3. Next, we use Figure 6 to illustrate some of our findings.

Figure 6 shows a sample path of our trace-driven simulations. We let  $\alpha = 1$ ,  $M = 30$ , which corresponds to 2:30 hours (see Figure 2(a)),  $\tau = 100$ , which corresponds to up to 3 updates per day,  $p = 0.54$  (see Figure 2(b)),  $G = 0.4$ ,  $P = \hat{B} = 40$  and  $T = 11$ . The number of users is initially 50, and decreases to 20 at round 200. We assume that users solve the USER PROBLEM when setting their activation strategies, and Algorithm 1 is run by the publisher every 100 rounds. We consider half of the population in one bus and the other half in another (see Appendix D.3). Despite the correlations among users, Figure 6 does not qualitatively deviate from the results obtained with uncorrelated users. In particular, the algorithm converges in up to 20 rounds, and the number of transmissions per slot varies between 9 and 11. The oscillations are not always centered at  $T = 11$  due to the fact that the threshold adopted by the users is integer valued, which might prevent inequality (16) to bind. Note that the bonus level converges to values in the optimal range, obtained using Proposition 6.1 and marked with dotted lines in Figure 6. When  $N = 20$ , the publisher sponsors the service provider costs ( $B = P$ ), and users make their updates for free.

## 7 Discussion of Assumptions, Limitations and Future Work

Next, we discuss the main simplifications adopted to yield a tractable model (Assumption 3.1 is discussed in details in §4, so we do not include it in the following list).

**Frequent decisions assumption:** We assume that users are interested in maximizing their expected average rewards. This assumption is appropriate if decisions are made frequently. In our measurement study,

we assumed that users make decisions every 5 minutes, and we observed that for a vast number of bus shifts the system parameters are stable over hours.

**High load assumption:** We assume that updates are issued every time slot with high probability. As discussed in §2.3.1, this assumption holds for certain news feeds categories such as *top stories*. The high load assumption also holds if users synchronize the updates of multiple applications, using APIs such as OVI Notifications [15]. In this case, the larger the number of applications that require updates, the higher the chances of at least one update being available every time slot. If multiple application updates can be performed at a single time slot, the bundle being subject to no additional costs or bonus, our model holds without modifications. However, if the interaction among the applications is non trivial, new policies need to be devised accounting for decisions such as when and from whom to download updates from each of the applications (see §8 for complementary work on that topic).

**Self-regarding users assumption:** We assume that users do not collaborate with each other. Although researchers are interested in leveraging collaboration among users [21, 22], a large number of mobile systems still does not take advantage of peer-to-peer transfers, and users need to download their messages exclusively from access points or base stations [23]. Our model applies to such systems. Nevertheless, we envision that if peers are roughly uniformly distributed, peer-assisted opportunistic transfers can be easily captured by our model simply by adapting the contact probability,  $p$ , in order to account for peer-to-peer contacts. Future work consists of accounting for spatial information when modeling the likelihood of contact opportunities.

Finally, note that in this paper we do not consider prediction algorithms in order to infer future opportunities with WiFi access points. In particular, some of the algorithms without prediction previously proposed in the literature, such as [3, §6.1.1], are special instances of the aging control policies described in this paper. As shown in §4, neglecting the possibility of prediction comes with no significant loss in scenarios of practical interest, the incorporation of prediction algorithms in our framework being subject for future work.

## 8 Related Work

The literature on measuring [19, 24, 23], modeling [25, 21, 26, 27, 28] and control [29, 30, 31, 22, 32, 33, 34] in wireless networks is vast. Nevertheless, we were not able to find any previous study on the *aging control* problem as described in this paper. Previous work accounted for the modeling of aging [25] or for the age control by publishers [29, 31], but not for users aging control as described in this paper. We were also not able to find any previous study on the the implications of *bonus packages* set by content and/or service providers [5].

Chaintreau *et al.* [25] model the distribution of message ages in a large scale mobile network using a spatial mean field approach. Their model allows the analysis of gossiping through opportunistic contacts. In this paper, in contrast, we assume that nodes rely exclusively on base stations and access points in order to receive their updates.

Activation control strategies were proposed in [35, 29, 30, 11]. In [29] the authors consider publishers of

evolving files, that aim at reducing their energy expenditure by controlling the probability of transmitting messages to users. In [30], a joint activation and transmission control policy is proposed so as to maximize the throughput of users under energy constraints. In [11], a joint activation and link selection control policy is proposed so as to minimize the energy consumption under delay constraints. Our work differs from [29, 30, 11] in two ways as 1) we investigate the activation policy of mobile nodes based on the utility of their messages and 2) we study the publishers bonus strategy. In addition, our analysis is carried out using MDPs and is validated using UMass Dieselnets traces (similar methodology is adopted, for instance, by Yang *et al.* [36]).

The utility function introduced in this paper corresponds to the impatience function presented by Reich and Chaintreau [31]. Reich and Chaintreau [31] study the implications of delays between requests and services on optimal content replication schemes. If users have limited caches and cannot download all the requested files every time they are in range of an access point, the insights provided by [31, 37] need to be coupled with the ones presented here in order to devise the optimal joint activation-replication strategy. Therefore, aging control, as described in this paper, significantly complements replication control, as described in [31, 37].

## 9 Conclusion

This paper reports our measurement and modeling studies of aging control in hybrid wireless networks. From the DieselNet measurements, we learned that correlations among contact opportunities do not play a key role if the energy costs incurred by users are small or if users cannot discriminate their strategies based on bus shifts. We then modeled and solved the aging control problem, and used trace-driven simulations to show that a very simple threshold strategy derived from our model performs pretty well in practice. When publishers are strategic, we analyzed the bonus package selection problem, and showed that in some scenarios it is beneficial for publishers to fully sponsor the content updates requested by users. We believe that the study of mechanisms to support applications that require frequent updates, such as microblogging, in wireless networks, is an interesting field of research, and we see our paper as a first attempt to shed light into the tradeoffs faced by users and publishers of such applications.

## References

- [1] Twitter Inc., “Twitter,” 2010, <http://twitter.com>.
- [2] “Exchange activesync,” 2011, <http://www.microsoft.com/>.
- [3] A. Balasubramanian, R. Mahajan, and A. Venkataramani, “Augmenting mobile 3g using wifi,” in *MobiSys*, 2010.
- [4] N. Banerjee, M. Corner, and B. Levine, “Energy efficient architecture for dtn throwboxes,” in *Infocom*, 2007.
- [5] SFR, “Bénéficiez de Bonus Fidélité,” 2010, <http://www.lepactesfr.fr/>.
- [6] R. T. Ma, D. Chiu, J. C. Lui, V. Misra, and D. Rubenstein, “On cooperative settlement between content, transit and eyeball internet service providers,” in *CoNEXT*, 2008.
- [7] J. Wortham, “Customers angered as iphones overload 3g,” 2009, <http://nytimes.com/>.
- [8] “T-Mobile @ Home,” 2009, <http://t-mobile.com/>.

- [9] N. Balasubramanian, A. Balasubramanian, and A. Venkataramani, "Energy consumption in mobile phones," in *IMC*, 2009.
- [10] "Text messaging," 2011, [http://en.wikipedia.org/wiki/Text\\_messaging](http://en.wikipedia.org/wiki/Text_messaging).
- [11] M.-R. Ra, J. Paek, A. B. Sharma, R. Govindan, M. H. Krieger, and M. J. Neely, "Energy-delay tradeoffs in smartphone applications," in *MobiSys*, 2010.
- [12] "Free wifi hotspots in Paris," 2011, <http://goparis.about.com/>.
- [13] "Amherst's community wireless mesh network with free internet access," 2011, <http://www.amherstma.gov/>.
- [14] "The new data plans of tim," 2011, <http://www.gizmodo.com.br/>.
- [15] "Ovi notifications api 2013 intelligent way to stay connected to the network," 2011, <http://blogs.forum.nokia.com>.
- [16] "Juice defender," 2011, <http://latedroid.com/juicedefender>.
- [17] "Free wifi manager," 2011, <http://www.freewifimanager.fr/>.
- [18] M. L. Puterman, *Markov Decision Processes*. Academic Press, 1994.
- [19] A. Balasubramanian, B. N. Levine, and A. Venkataramani, "Enabling Interactive Applications for Hybrid Networks," in *MOBICOM*, 2008.
- [20] H. J. Kushner and G. G. Yin, *Stochastic Approximation and Recursive Algorithms and Applications*. Springer, 2003.
- [21] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, and C. Diot, "Pocket switched networks," in *UCAM-CL-TR-617*, 2005.
- [22] E. Altman, G. Neglia, F. D. Pellegrini, and D. Miorandi, "Decentralized stochastic control of delay tolerant networks," in *INFOCOM*, 2009.
- [23] D. Gunawardena, T. Karagiannis, A. Proutiere, and M. Vojnovic, "Characterizing podcast services," in *IMC*, 2009.
- [24] X. Zhang, J. Kurose, B. N. Levine, D. Towsley, and H. Zhang, "Study of a Bus-Based Delay Tolerant Network," in *MOBICOM*, 2007.
- [25] A. Chaintreau, J.-Y. L. Boudec, and N. Ristanovic, "The age of gossip: spatial mean field regime," in *SIGMETRICS*, 2009.
- [26] K.-W. Kwong, A. Chaintreau, and R. Guerin, "Quantifying content consistency improvements through opportunistic contacts," in *4th ACM Workshop on Challenged Networks*, 2009.
- [27] R. Bakhshi, J. Endrullis, S. Endrullis, W. Fokkink, and B. Haverkort, "Automating the mean-field method for large dynamic gossip networks," in *QEST*, 2010.
- [28] N. Banerjee, M. D. Corner, D. Towsley, and B. N. Levine, "Relays, base stations, and meshes: enhancing mobile networks with infrastructure," in *MOBICOM*, 2008.
- [29] E. Altman, P. Nain, and J. Bermond, "Distributed storage management of evolving files in delay tolerant ad hoc networks," in *INFOCOM*, 2009.
- [30] E. Altman, A. P. Azad, T. Basar, and F. D. Pellegrini, "Optimal Activation and Transmission Control in DTNs," in *INFOCOM*, 2009.
- [31] J. Reich and A. Chaintreau, "The age of impatience: optimal replication schemes for opportunistic networks," in *CONEXT*, 2009.
- [32] I. Menache and E. Altman, "Battery-state dependent power control as a dynamic game," in *WiOpt*, 2009.
- [33] Q. Yuan, I. Cardei, and J. Wu, "Predict and relay: An efficient routing in disruption-tolerant networks," in *Mobihoc*, 2009.
- [34] C. Liu and J. Wu, "An optimal probabilistic forwarding protocol in delay tolerant networks," in *Mobihoc*, 2009.
- [35] L. B. Le, E. Modiano, and N. B. Shroff, "Optimal control of wireless networks with finite buffers," in *Infocom*, 2010.
- [36] R. Yang, S. Bhulai, R. van der Mei, and F. Seinstra, "Optimal resources allocation for time reservation systems," *Performance Evaluation*, vol. 68, pp. 414–428, 2011.

- [37] S. Ioannidis, A. Chaintreau, and L. Massoulie, “Optimal and scalable distribution of content updates over a mobile social network,” in *INFOCOM*, 2009.
- [38] R. Cavazos-Cadena, “Necessary conditions for the optimality equation in average-reward markov decision processes,” *Applied Mathematical Optimization*, vol. 19, pp. 97–112, 1989.
- [39] —, “Necessary and sufficient conditions for a bounded solution to the optimality equation in average reward markov decision chains,” *Systems and Control Letters*, vol. 10, pp. 71–78, 1988.
- [40] V. Borkar, *Stochastic approximation: A Dynamic Systems Viewpoint*. Cambridge University Press, 2008.
- [41] A. F. Filippov, *Differential equations with discontinuous righthand sides*. Kluwer Academic Publishers, 1988.
- [42] D. Shevitz and B. Paden, “Lyapunov stability theory of nonsmooth systems,” *Trans. on Auto. Cont.*, vol. 39, no. 9, 1994.
- [43] P. Key, L. Massoulie, and B. Wang, “Emulating low-priority transport at the application layer: a background transfer service,” in *ACM Sigmetrics*, 2004.



## A Derivation of Main Results When $P_{3G} \rightarrow \infty$

We begin analyzing the case in which 3G is not available. The optimality conditions are

$$\boxed{V(x) + g_u = \max \left( U(x) + V(\min(x+1, M)), U(x) - G + p(V(1) - P + B) + (1-p)V(\min(x+1, M)) \right), \quad x=1, \dots, M} \quad (17)$$

Let

$$H(x, 0) = U(x) + V(\min(x+1, M)) \quad (18)$$

$$H(x, 1) = U(x) - G + p(V(1) - P + B) + (1-p)V(\min(x+1, M)) \quad (19)$$

It follows from (17)-(19) and [18] that the following policy is optimal, for  $m = 1, \dots, M$ ,

$$a(m) = \begin{cases} 0, & \text{if } H(m, 0) \geq H(m, 1) \\ 1, & \text{otherwise} \end{cases} \quad (20)$$

At states  $M$  and  $M-1$ , (17) implies that

$$\begin{cases} (H(M, 0) \geq H(M, 1)) \wedge (H(M-1, 0) \geq H(M-1, 1)), & \text{if } V(M) \geq U(M) - G + p(V(1) - P + B) + (1-p)V(M) \\ (H(M, 0) \leq H(M, 1)) \wedge (H(M-1, 0) \leq H(M-1, 1)), & \text{otherwise} \end{cases} \quad (21)$$

Equation (21) yields the following remark, used in the analysis of the base cases of the inductive arguments that follow.

**Remark A.1.**  $(H(M, 0) \geq H(M, 1)) \iff (H(M-1, 0) \geq H(M-1, 1)).$

### A.1 Derivation of (12) and corresponding steady state probabilities

From §5.2,

$$\pi_1 = \pi_i, \quad i = 1, \dots, s \quad (22)$$

$$\pi_i = \pi_{i-1}(1-p) = \pi_1(1-p)^{i-s}, \quad i = s+1, \dots, M-1 \quad (23)$$

$$\pi_M = (1-p)(\pi_{M-1} + \pi_M) = \frac{1-p}{p}\pi_{M-1} \quad (24)$$

Therefore,

$$\pi_1 = \left( s + \sum_{i=s+1}^{M-1} (1-p)^{i-s} + (1-p)^{M-s}/p \right)^{-1} = \left( s + \sum_{i=1}^{M-1-s} (1-p)^i + (1-p)^{M-s}/p \right)^{-1} = \left( s + \frac{1-p}{p} \right)^{-1}$$

## A.2 Derivation of expression of $E[r; s]$

Next, we show that

$$E[r; s] = \pi_1 \left[ \sum_{x=1}^{s-1} U(x) + \sum_{i=0}^{M-1-s} U(i+s)(1-p)^i - \frac{G}{p} - P + B \right] \quad (25)$$

Let

$$\gamma = \left( (U(M) - G - pP + pB)\pi_M - \sum_{i=s}^{M-1} (G + pP - pB)\pi_i \right) / \pi_1 \quad (26)$$

Then,

$$E[r; s] = \sum_{i=1}^{s-1} U(i)\pi_1 + \sum_{i=s}^M (U(i) - G - pP + pB)\pi_i = \pi_1 \left[ \sum_{i=1}^{s-1} U(i) + \sum_{i=s}^{M-1} U(i)\pi_i/\pi_1 + \gamma \right] \quad (27)$$

$$\begin{aligned} &= \pi_1 \left[ \sum_{i=1}^{s-1} U(i) + \sum_{i=0}^{M-1-s} U(i+s)(1-p)^i + \frac{1}{p}(1-p)^{M-s}(U(M) - G - pP + pB) - (G + pP - pB) \frac{1 - (1-p)^{M-s}}{p} \right] \\ &= \pi_1 \left[ \sum_{i=1}^{s-1} U(i) + \sum_{i=0}^{M-1-s} U(i+s)(1-p)^i - G/p - P + B \right] \end{aligned} \quad (28)$$

## A.3 Proof of Proposition A.1

**Proposition A.1.** *The USER PROBLEM admits an optimal threshold policy*

$$a(x) = \begin{cases} 0, & \text{if } x < s^* \\ 1, & \text{if } s^* \leq x \leq M. \end{cases} \quad (29)$$

*Proof.* We show that, for  $m = 0, \dots, M-1$ ,

•

$$V(M-m-1) - V(M-m) \geq 0 \quad (30)$$

•

$$(H(M-m, 0) \geq H(M-m, 1)) \Rightarrow (H(M-m-1, 0) \geq H(M-m-1, 1)) \quad (31)$$

We consider two scenarios,  $H(M, 0) \geq H(M, 1)$  and  $H(M, 0) < H(M, 1)$ .

$$\boxed{\boxed{\text{scenario 1) } H(M, 0) \geq H(M, 1)}} \quad (32)$$

If  $H$  satisfies (31) then, from (17)-(19),

$$V(m) = U(m) + V(m+1), 1 \leq m \leq M-1 \quad (33)$$

and (30) holds.

Next, assume for the sake of contradiction that  $H$  does not satisfy (31). Let  $m$  be the largest state at which condition (31) is violated,

$$m = \max\{i | H(M-i, 0) \geq H(M-i, 1) \text{ and } H(M-i-1, 1) > H(M-i-1, 0)\} \quad (34)$$

It follows from (34) that

$$H(M-m-1, 1) > H(M-m-1, 0) \quad (35)$$

$$H(M-m+k, 0) \geq H(M-m+k, 1), k = 0, 1, \dots, m \quad (36)$$

(35) and (36) yield, respectively,

$$V(M-m) - V(1) < -G/p - P + B \quad (37)$$

$$V(M-k) - V(1) \geq -G/p - P + B, \quad k = 0, \dots, m-1 \quad (38)$$

Letting  $k = m-1$  in (38),

$$V(M-m) = U(M-m) + V(M-m+1) \geq U(M-m) + V(1) - G/p - P + B \quad (39)$$

(37) and (39) yield the following contradiction

$$V(1) - G/p - P + B \leq V(M-m) < V(1) - G/p - P + B \quad (40)$$

Therefore, (31) holds for  $m = 0, \dots, M-1$ .

scenario 2)  $H(M, 0) < H(M, 1)$

(41)

**Base case:** We first show that  $V(M-1) \geq V(M)$ .

Note that  $(H(M, 1) \geq H(M, 0)) \Rightarrow (H(M-1, 1) \geq H(M-1, 0))$  (see remark A.1). It follows from (17) that

$$V(M-1) = U(M-1) - G - P + B + pV(1) + (1-p)V(M) - g_u \quad (42)$$

$$V(M) = U(M) - G - P + B + pV(1) + (1-p)V(M) - g_u \quad (43)$$

Hence, (42)-(43), together with the fact that  $U(x)$  is non-increasing, yield  $V(M-1) \geq V(M)$ .

**Induction hypothesis [assume result holds for  $m < t$ ]:** Assume that  $V(M-m-1) - V(M-m) \geq 0$ ,

for  $m < t$ , and  $(H(M - m, 0) \geq H(M - m, 1)) \Rightarrow (H(M - m - 1, 0) \geq H(M - m - 1, 1))$ , for  $m < t$ .

**Induction step [show result holds for  $m = t$ ]:** Next, we show that  $V(M - t - 1) - V(M - t) \geq 0$  and that  $(H(M - t, 0) \geq H(M - t, 1)) \Rightarrow (H(M - t - 1, 0) \geq H(M - t - 1, 1))$ .

It follows from the induction hypothesis that  $V(M - t + 1) \leq V(M - t)$ . We consider two cases,

i)  $H(M - t, 0) \geq H(M - t, 1)$  The proof is similar to that of scenario 1.

ii)  $H(M - t, 1) \geq H(M - t, 0)$  Now, we show that

$$V(M - t - 1) - V(M - t) \geq 0.$$

Note that  $(H(M - t, 0) \geq H(M - t, 1)) \Rightarrow (H(M - t - 1, 0) \geq H(M - t - 1, 1))$  holds vacuously. If  $(H(M - t, 1) \geq H(M - t, 0))$  and  $(H(M - t - 1, 1) \geq H(M - t - 1, 0))$ ,

$$V(M - t) + g_u = U(M - t) - G - P + B + pV(1) + (1 - p)V(\min(M - t + 1, M)) \quad (44)$$

$$V(M - t - 1) + g_u = U(M - t - 1) - G - P + B + pV(1) + (1 - p)V(\min(M - t, M)) \quad (45)$$

Also,  $V(M - t + 1) \leq V(M - t)$  (induction hypothesis) and  $U(M - t) < U(M - t - 1)$  (by assumption). Hence, (44)-(45) yield  $V(M - t) \leq V(M - t - 1)$ .

□

**Remark A.2.** *In what follows our analysis is restricted to threshold policies.*

## A.4 On the number of optimal thresholds

The existence of policies satisfying (17) follows from [38] and [39]. In particular, the optimal threshold policies satisfy (17). In general, though, the solution to (17) is not unique. Next, we characterize the scenarios in which there are two or more optimal threshold policies.

Figure 7 illustrates scenarios in which Assumption A.1 does not hold. Let  $b = 0.3$ ,  $M = 21$ ,  $p = 0.5$  and  $U(x) = c$  if  $x \leq 3$  and  $U(x) = 0$  otherwise. Figures 7(b), 7(d) and 7(f) show the three utility functions considered, corresponding to  $c = 4, 12$  and  $16$ , respectively. If  $c = 4$ , threshold policies with  $s \geq 3$  yield optimal reward of zero. If  $c = 12$ , there are two optimal thresholds ( $s = 2, 3$ ). Finally, if  $c = 16$  the optimal threshold policy is unique ( $s = 2$ ).

Proposition A.2 establishes conditions according to which the number of optimal thresholds is at most two.

**Proposition A.2.** *Let  $R$  be the number of optimal thresholds.*

$$\begin{cases} R \geq 3, & \text{if } \exists m < M-1 \text{ s.t. } \sum_{x=1}^{m-1} U(x) = G/p + P - B \text{ and } U(x) = 0 \quad \forall x > m \\ R \leq 2, & \text{otherwise} \end{cases} \quad (46)$$

*If  $R \geq 3$  then the policy always inactive is optimal.*

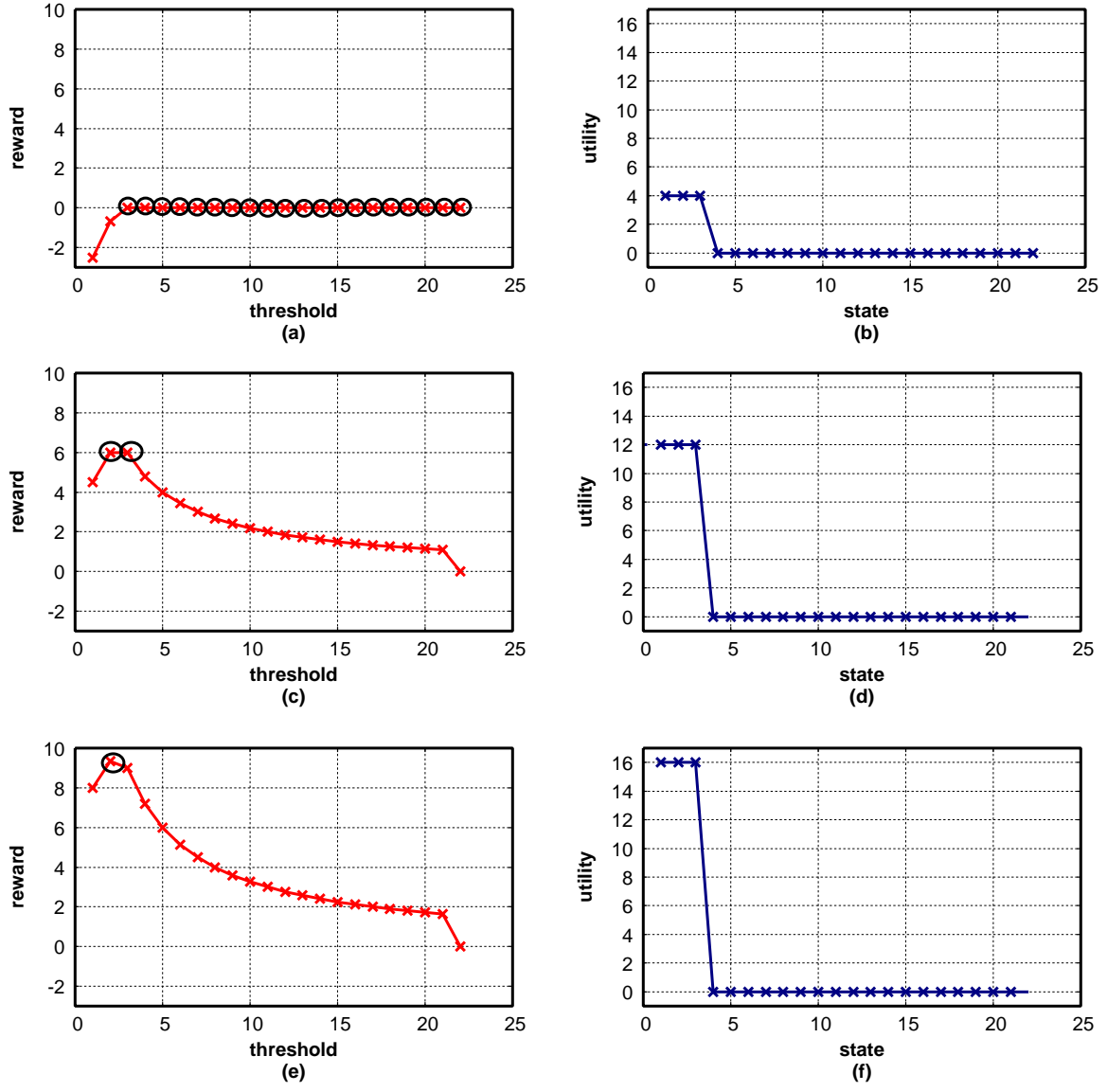


Figure 7: Illustration of scenarios in which the optimal threshold is not unique ((a) and (c)) and unique (e).

*Proof.* From proposition A.3, the expected reward  $E[r; s]$  of a threshold policy is non-decreasing when  $s \leq s^*$ , and is non-increasing when  $s \geq s^*$ . Therefore, the optimal thresholds should be consecutive. Let  $s^* - 1$ ,  $s^*$  and  $s^* + 1$  be three optimal thresholds,

$$E[r; s^* - 1] = E[r; s^*] = E[r; s^* + 1] \quad (47)$$

From (47) and (69),

$$\left(s^* - 1 + \frac{1-p}{p}\right) E[r; s^* - 1] - \left(s^* + \frac{1-p}{p}\right) E[r; s^*] = -E[r; s^*] = -p \sum_{i=1}^{M-s^*} U(i + s^* - 1)(1-p)^i \quad (48)$$

$$\left(s + \frac{1-p}{p}\right) E[r; s^*] - \left(s^* + 1 + \frac{1-p}{p}\right) E[r; s^* + 1] = -E[r; s^*] = -p \sum_{i=1}^{M-s^*-1} U(i + s^*)(1-p)^{i-1} \quad (49)$$

Subtracting (49) from (48),

$$U(s^*) = 0 \quad (50)$$

The above equation together with the fact that  $U(x)$  is a non-increasing function yields

$$U(x) = 0, \quad x \geq s^* \quad (51)$$

(51) together with (48) and (49) yields

$$E[r; s^* - 1] = E[r; s^*] = E[r; s^* + 1] = E[r; M + 1] = 0 \quad (52)$$

Substituting (51) and (52) into (25) yields

$$\sum_{x=1}^{s^*-1} U(x) - \frac{G}{p} - P + B = 0. \quad (53)$$

Therefore, if (53) and (51) hold then  $R \geq 3$ . Otherwise,  $R \leq 2$ .

□

**Corollary A.1.** *If the policy always inactive is not optimal, the number of optimal thresholds is at most two.*

**Corollary A.2.** *If the optimal reward is greater than zero the number of optimal thresholds is at most two.*

**Corollary A.3.** *Let  $s^*$  be an optimal threshold. If  $E[r; s^*] \neq E[r; s^* + 1]$  and  $E[r; s^*] \neq E[r; s^* - 1]$  then the optimal threshold is unique.*

Given the above characterization of the cases in which the optimal threshold policy is not unique, in the rest of this Appendix we consider the following assumption.

**Assumption A.1.** *There is at most one optimal threshold policy.*

## A.5 Proof of Proposition 5.1

*Proof.* Let  $a$  be the optimal threshold policy.

### A.5.1 Conditions for the optimal policy to be always inactive

It follows from the optimality conditions (17) and Assumption A.1 that

$$a(x) = 0, 1 \leq x \leq M \iff H(x, 1) \leq H(x, 0), 1 \leq x \leq M \quad (54)$$

Therefore, from the above equation and (18)-(19),

$$a(x) = 0, 1 \leq x \leq M \iff V(1) - V(x) \leq G/p + P - B, 1 \leq x \leq M \quad (55)$$

Since  $V(x)$  is decreasing (see Proposition A.1), (55) yields

$$a(x) = 0, 1 \leq x \leq M \iff V(1) - V(M) \leq G/p + P - B \quad (56)$$

or equivalently,

$$a(x) = 0, 1 \leq x \leq M \iff V(M) \geq V(1) - G/p - P + B \quad (57)$$

From (17),

$$a(x) = 0, 1 \leq x \leq M \iff V(i) = U(i) + V(i+1), 1 \leq i \leq M-1 \quad (58)$$

where  $\Leftarrow$  in (58) follows from the assumption that there is a unique optimal policy that satisfies (17).

Therefore,

$$a(x) = 0, 1 \leq x \leq M \iff V(j) = \sum_{i=j}^M U(i), j = 1, \dots, M \quad (59)$$

Finally, (57) and (59) yield

$$a(x) = 0, 1 \leq x \leq M \iff 0 \geq \sum_{i=1}^M U(i) - G/p - P + B \quad (60)$$

### A.5.2 Conditions for the optimal policy to be always active

It follows from the optimality conditions (17) and Assumption A.1 that

$$a(x) = 1, 1 \leq x \leq M \iff H(x, 1) \geq H(x, 0), 1 \leq x \leq M \quad (61)$$

Therefore, from the above equation and (18)-(19),

$$a(x) = 1, 1 \leq x \leq M \iff V(1) - V(x) \geq G/p + P - B, 1 \leq x \leq M \quad (62)$$

Since  $V(x)$  is decreasing (see Proposition A.1), it follows from (62) that

$$a(x) = 1, 1 \leq x \leq M \iff V(1) - V(2) \geq G/p + P - B \quad (63)$$

or equivalently,

$$a(x) = 1, 1 \leq x \leq M \iff V(2) < V(1) - G/p - P + B \quad (64)$$

Letting  $x = 1$  in (17) yields

$$a(x) = 1, 1 \leq x \leq M \iff V(1) = -G + U(1) + pV(1) - pP + pB + (1-p)V(2) - E[r; 1] \quad (65)$$

where  $\Leftarrow$  in (65) follows since our analysis is restricted to threshold policies. Note that  $a(1) = 1$  is implied by the right hand side of (65) together with (17). If  $a(1) = 1$  then  $a(x) = 1, 1 \leq x \leq M$ .

Finally, (64) and (65) yield

$$V(1) = -G + U(1) + pV(1) - pP + pB + (1-p)V(2) - E[r; 1] \iff \quad (66)$$

$$\iff V(1) \leq -G + U(1) + pV(1) - pP + pB + (1-p)(V(1) - G/p - P + B) - E[r; 1] \quad (67)$$

The desired result follows from algebraic manipulation of (67),

$$a(x) = 1, 1 \leq x \leq M \iff E[r; 1] < U(1) - G/p - (P - B) \quad (68)$$

□

**Remark A.3.** If Assumption A.1 does not hold, the proof of Proposition 5.1 presented above remains valid after replacing all  $\iff$  by  $\Rightarrow$ .

## A.6 Proof of Proposition 5.2

The proof of Proposition 5.2 follows directly from Proposition A.3.

**Proposition A.3.** If  $s \leq s^* - 1$  then  $E[r; s] \leq E[r; s + 1]$ . Otherwise,  $E[r; s] \geq E[r; s + 1]$ .

*Proof.* Next, we show that if  $s \leq s^* - 1$  then  $E[r; s] \leq E[r; s + 1]$  (the other case follows similarly). Let  $s^*$  be the optimal threshold. It follows from Definition 3.1 that  $E[r; s] - E[r; s^*] \leq 0$  for  $1 \leq s \leq M$ . Next, we show that  $E[r; s^* - m] \geq E[r; s^* - (m + 1)]$  for  $m < s^*$ . The proof is by induction on  $m$ .



Algebraic manipulation of (25) yields

$$\begin{aligned}
& \left( s - 1 + \frac{1-p}{p} \right) E[r; s-1] - \left( s + \frac{1-p}{p} \right) E[r; s] \\
= & \left[ \sum_{i=1}^{s-2} U(i) + \sum_{i=0}^{M-s} U(i+s-1)(1-p)^i - \frac{G}{p} - P + B \right] - \left[ \sum_{i=1}^{s-1} U(i) + \sum_{i=1}^{M-s} U(i+s-1)(1-p)^{i-1} - \frac{G}{p} - P + B \right] \\
= & -p \sum_{i=1}^{M-s} U(i+s-1)(1-p)^{i-1}. \tag{69}
\end{aligned}$$

Equation (69) yields

$$\left( s + \frac{1-p}{p} \right) (E[r; s-1] - E[r; s]) = E[r; s] - p \sum_{i=1}^{M-s} U(i+s)(1-p)^{i-1} \tag{70}$$

**Base case:** It follows from Definition 3.1 that the statement holds for  $m = 0$ . Note that  $E[r; s^* - 1] - E[r; s^*] \leq 0$  yields

$$\left( s^* - 1 + \frac{1-p}{p} \right) (E[r; s^* - 1] - E[r; s^*]) \tag{71}$$

$$\stackrel{(*)}{=} E[r; s^*] - p \sum_{i=1}^{M-s^*} U(i+s^*-1)(1-p)^{i-1} \tag{72}$$

$$= \left( s^* + \frac{1-p}{p} \right) (E[r; s^* - 1] - E[r; s^*]) \tag{73}$$

$$\stackrel{(**)}{=} E[r; s^* - 1] - p \sum_{i=1}^{M-s^*} U(i+s^*-1)(1-p)^{i-1} \leq 0$$

$$\therefore E[r; s^* - 1] \leq p \sum_{i=1}^{M-s^*} U(i+s^*-1)(1-p)^{i-1} \tag{74}$$

where  $(*)$  follows from (25) and  $(**)$  follows from  $(*)$  after summing  $E[r; s^* - 1] - E[r; s^*]$  to both sides.

**Induction hypothesis:** Assume that  $E[r; s^* - m] \geq E[r; s^* - (m+1)]$  for  $m < t$ .

**Induction step:** We show that the proposition holds for  $m = t$ . To this goal, we compare  $E[r; s^* - t]$

and  $E[r; s^* - (t + 1)]$ ,

$$\begin{aligned}
& \left( s^* - t + \frac{1-p}{p} \right) (E[r; s^* - t] - E[r; s^* - t - 1]) \\
& \stackrel{(*)}{=} p \sum_{i=1}^{M-s^*+t} U(i + s^* - t - 1)(1-p)^{i-1} - E[r; s^* - t] \\
& > p \sum_{i=1}^{M-s^*} U(i + s^* - t - 1)(1-p)^{i-1} - E[r; s^* - t] \\
& > p \sum_{i=1}^{M-s^*} U(i + s^*)(1-p)^{i-1} - E[r; s^* - t] \\
& \stackrel{(**)}{\geq} p \sum_{i=1}^{M-s^*} U(i + s^*)(1-p)^{i-1} - E[r; s^* - 1] \stackrel{(***)}{\geq} 0
\end{aligned}$$

where  $(*)$  follows from (70),  $(**)$  follows from the induction hypothesis and  $(***)$  follows from (74). The proof is completed by noting that  $s^* - t + \frac{1-p}{p} \geq 0$  hence  $E[r; s^* - t] \geq E[r; s^* - t - 1]$ .  $\square$

## A.7 Proof of Proposition 5.3

*Proof.* In what follows we show that the optimal threshold increases with respect to  $G$ . The proof that the optimal threshold increases with respect to  $P$  and decreases with respect to  $B$  is similar. From (28),

$$\frac{d}{dG} E[r; s] = -\frac{1}{ps + 1 - p} \quad (75)$$

Let  $s_1 > s_0$ . Then,

$$\lim_{\Delta G \rightarrow 0} \frac{E[r; s_1, G + \Delta G] - E[r; s_1, G]}{\Delta G} > \lim_{\Delta G \rightarrow 0} \frac{E[r; s_0, G + \Delta G] - E[r; s_0, G]}{\Delta G} \quad (76)$$

From (28), it also follows that

$$E[r; s, G + \Delta G] - E[r; s, G] = -\frac{1}{s + (1-p)/p} \frac{\Delta G}{p} \quad (77)$$

Therefore,

$$E[r; s_1, G] - E[r; s_1, G + \Delta G] < E[r; s_0, G] - E[r; s_0, G + \Delta G] \quad (78)$$

Assume, for the sake of contradiction, that  $s_1$  and  $s_0$  are optimal thresholds when the energy cost is  $G$  and  $G + \Delta$ , respectively,

$$E[r; s_1, G] \geq E[r; s, G], \quad s \neq s_1 \quad (79)$$

$$E[r; s_0, G + \Delta G] \leq E[r; s, G + \Delta G], \quad s \neq s_0 \quad (80)$$

In particular,

$$E[r; s_1, G] \geq E[r; s_0, G] \quad (81)$$

$$E[r; s_0, G + \Delta G] \leq E[r; s_1, G + \Delta G] \quad (82)$$

Then,  $E[r; s_1, G] - E[r; s_1, G + \Delta G] \geq E[r; s_0, G] - E[r; s_0, G + \Delta G]$ , which contradicts the (78).  $\square$

**Remark A.4.** In (78),  $<$  should be replaced by  $\leq$  if Assumption A.1 does not hold. In this case, the optimal threshold is non decreasing with respect to  $G$ .

## A.8 Proof of Proposition 5.4 (step utility)

*Proof.* We assume that the optimal policy consists of remaining active in at least one state (otherwise,  $s^* = M + 1$  and  $E[r; s] = 0$ ). We consider two cases,

i)  $k > s$ :

$$E[r; s] = \frac{1}{s + \frac{1-p}{p}} \left[ vk - \frac{G}{p} - P + B \right] \quad (83)$$

ii)  $k \leq s$ :

$$E[r; s] = \frac{1}{s + \frac{1-p}{p}} \left[ vs + v \sum_{i=1}^{k-s} (1-p)^i - \frac{G}{p} - P \right] = \frac{1}{s + \frac{1-p}{p}} \left[ vs + \frac{v(1-p)}{p} - \frac{v(1-p)^{k-s+1}}{p} - \frac{G}{p} - P + B \right]$$

We now show that in the above two cases the optimal threshold can be efficiently computed by comparing the values of  $E[r; s]$  at five points. To this goal, we first assume that the optimal threshold,  $s^*$ , can take real values (the assumption will be removed in the next paragraph). We refer to  $s^*$  as an interior maximum if  $1 < s^* < M$ , and as a boundary maximum if  $s^* = 1$  or  $s^* = M$ . Then, a necessary condition for  $s^*$  to be an interior maximum of  $E[r; s]$  consists of  $s^*$  being a root of  $\frac{d}{ds} E[r; s] = 0$ . If  $k > s$  then  $\frac{d}{ds} E[r; s] = 0$  has no roots in the interval  $[1, M]$ , since (83) is monotonic with respect to  $s$ . If  $k \leq s$ , let  $\varphi$  be the root of  $\frac{d}{ds} E[r; s] = 0$ . Then,

$$\varphi = - \left( p + W \left( (G + pP - pB) \exp \left( - \frac{\ln(1-p)(1+kp) + p}{p} \right) v^{-1} \right) p + \ln(1-p)(1-p) \right) / (\ln(1-p)) \quad (84)$$

In (84),  $W(x)$  denotes the Lambert function, i.e.,  $W(x) = w$  if  $we^w = x$ .

Accounting for the fact that  $s^*$  is an integer, and that it might be either an interior or boundary maximum, yields,

$$s^* = \begin{cases} 1, & \text{if } k > s^* \text{ and } vk - G/p - P + B > 0 \\ k - 1, & \text{if } k > s^* \text{ and } vk - G/p - P + B \leq 0 \\ \psi = \min(\lceil \varphi \rceil, M), & \text{if } k \leq s^* \text{ and } E[r; \psi] \geq E[r; \psi + 1] \\ \min(\lfloor \varphi \rfloor, M), & \text{otherwise} \end{cases} \quad (85)$$

□

Therefore,  $s^*$  can be computed comparing the value of  $E[r; s]$  at  $s = 1, k - 1, \lfloor \varphi \rfloor, \lceil \varphi \rceil, M$ .

## A.9 Linear utility

When  $U(x) = M - x$ , the expected reward is

$$E[r; s] = \pi_1 \left[ (s-1)M - s(s-1)/2 + \frac{(1-p)^{M-s+1} + pM - ps - 1 + p}{p^2} - \frac{G}{p} - P + B \right]$$

## A.10 Expected age

If  $p > 0$ , the expected age,  $A$ , is  $A = \sum_{i=1}^M i\pi_i$ ,

$$\begin{aligned} A &= \pi_1 \left[ \sum_{i=1}^s i + \sum_{i=s+1}^{M-1} i(1-p)^{i-s} + \frac{(1-p)^{M-s}}{p} M \right] \\ &= \sum_{i=1}^M i\pi_i = M \frac{1-p}{p} (1-p)^{M-1-s} \frac{1}{s + (1-p)/p} + \sum_{i=1}^s i \frac{1}{s + (1-p)/p} + \sum_{i=s+1}^{M-1} (1-p)^{i-s} i \frac{1}{s + (1-p)/p} \\ &= \frac{p^2 s^2 - p^2 s - 2(1-p)^{M-s}(1-p) + 2sp + 2 - 2p}{2p(sp + 1 - p)} \end{aligned} \quad (86)$$

## A.11 Proposition A.4

**Proposition A.4.** *The expected age is an increasing function of  $s$ .*

*Proof.* Let  $A(s)$  be the average age of a threshold policy whose threshold is  $s$ . Consider two threshold policies, with corresponding thresholds  $s$  and  $s+1$ . Then,

$$A(s) = \sum_{i=1}^s \frac{i}{s + \frac{1-p}{p}} + \sum_{i=s+1}^{M-1} \frac{i(1-p)^{i-s}}{s + \frac{1-p}{p}} + \frac{(1-p)^{M-s}}{p} \frac{M}{s + \frac{1-p}{p}} \quad (87)$$

$$A(s+1) = \sum_{i=1}^{s+1} \frac{i}{s+1 + \frac{1-p}{p}} + \sum_{i=s+2}^{M-1} \frac{i(1-p)^{i-s-1}}{s+1 + \frac{1-p}{p}} + \frac{(1-p)^{M-s-1}}{p} \frac{M}{s+1 + \frac{1-p}{p}} \quad (88)$$

Algebraic manipulation of (87)-(88) yields

$$A(s) \frac{s + \frac{1-p}{p}}{s+1 + \frac{1-p}{p}} = \sum_{i=1}^s \frac{i}{s+1 + \frac{1-p}{p}} + \sum_{i=s+1}^{M-1} \frac{i(1-p)^{i-s}}{s+1 + \frac{1-p}{p}} + \frac{(1-p)^{M-s}}{p} \frac{M}{s+1 + \frac{1-p}{p}} \quad (89)$$

$$A(s+1)(1-p) = \sum_{i=1}^{s+1} \frac{i(1-p)}{s+1 + \frac{1-p}{p}} + \sum_{i=s+2}^{M-1} \frac{i(1-p)^{i-s}}{s+1 + \frac{1-p}{p}} + \frac{(1-p)^{M-s}}{p} \frac{M}{s+1 + \frac{1-p}{p}} \quad (90)$$

Subtracting the left-hand side of (90) from the left-hand side of (89), and denoting the difference by  $\Delta_1$ , yields

$$\begin{aligned}
\Delta_1 &= A(s) \frac{s + \frac{1-p}{p}}{s + 1 + \frac{1-p}{p}} - A(s+1)(1-p) \\
&= A(s) - A(s+1) - A(s) \frac{1}{s + 1 + \frac{1-p}{p}} + A(s+1)p \\
&= (A(s) - A(s+1))(1-p) + A(s) \frac{ps}{s + 1 + \frac{1-p}{p}}.
\end{aligned} \tag{91}$$

Subtracting the right-hand side of (90) from the right-hand side of (89), and denoting the difference by  $\Delta_2$ , yields

$$\Delta_2 = \sum_{i=1}^s \frac{pi}{s + 1 + \frac{1-p}{p}}. \tag{92}$$

Letting  $\Delta_1 = \Delta_2$ ,

$$(A(s) - A(s+1))(1-p) = \frac{ps}{s + 1 + \frac{1-p}{p}} \left( \frac{1}{s} \sum_{i=1}^s i - A(s) \right) \tag{93}$$

Next, we make the dependence of the age on  $p$  explicit, and denote it by  $A(s, p)$ .

The age is minimized when  $p = 1$ ,

$$A(s, 1) < A(s, p), \quad 0 < p < 1 \tag{94}$$

Letting  $p = 1$  in (86),

$$A(s, 1) = \frac{1}{s} \sum_{i=1}^s i. \tag{95}$$

(95) implies that the right hand side of (93) is negative. Therefore, it follows from (93)-(95) that  $A(s) < A(s+1)$  and, for  $s_1 < s_2$ ,  $A(s_1) < A(s_2)$ .  $\square$

## A.12 Proof of Proposition 6.1

*Proof.* Let  $f$  be a function that maps a bonus level into the corresponding optimal threshold selected by users that solve the USER PROBLEM,

$$f : B \rightarrow s^* \tag{96}$$

$$\mathbb{R} \rightarrow \mathbb{N} \tag{97}$$

Let  $g$  be a function that maps a threshold  $s$  into the *minimum* bonus level  $B$  such that  $s$  is the optimal

threshold under  $B$ ,

$$g : s \rightarrow B \quad (98)$$

$$\mathbb{N} \rightarrow \mathbb{R} \quad (99)$$

Let  $h$  be a function that maps a threshold  $s$  into the pair of bonus levels  $(\underline{B}(s), \overline{B}(s))$  such that  $s$  is the optimal threshold under any bonus in the range  $[\underline{B}(s), \overline{B}(s)]$ ,

$$h : s \rightarrow (\underline{B}(s), \overline{B}(s)) \quad (100)$$

$$\mathbb{N} \rightarrow \mathbb{R}^2 \quad (101)$$

Note that  $f$  admits no inverse (changing  $B$  might not alter  $s^*$ ). The function  $g$ , in contrast, is injective, since given  $B$  the optimal threshold  $s^*(B)$  is unique.

The publisher must choose  $B$  so as to minimize  $A$ . We consider two cases, varying according to the value of  $s(B)$  evaluated at  $B = P$ ,

case i)

$$s(P) < \frac{Np}{T} - \frac{1-p}{p}$$

In this case, inequality (16) is not satisfied when  $B = P$ . Since  $B$  cannot surpass  $P$ , the problem admits no solution.

case ii)

$$s(P) \geq \frac{Np}{T} - \frac{1-p}{p}$$

In this case, the problem admits a solution, to be derived next.

According to Proposition A.4, the average age  $A$  is an increasing function of  $s$ . According to Proposition 5.3, the optimal threshold  $s^*$  is a non-increasing function of  $B$ . Therefore, the objective function  $A$  of the PUBLISHER PROBLEM is a non-increasing function of  $B$ .

The fact that  $A$  is a non-increasing function of  $B$ , together with the fact that the l.h.s. of inequality (16) is increasing in  $B$ , imply that the optimal bonus consists of setting  $B$  as large as possible, without violating (16).

Inequality (16) binds when

$$s = \frac{Np}{T} - \frac{1-p}{p} \quad (102)$$

Therefore, the optimal threshold, which must be integer, is

$$s = \min \left( M + 1, \left\lceil \frac{Np}{T} - \frac{1-p}{p} \right\rceil \right) \quad (103)$$

The solution of the PUBLISHER PROBLEM consists of finding  $B$  such that  $s^* = s$ .

The bonus  $B$  must be chosen in the range  $[\underline{B}(s^*), \overline{B}(s^*)]$ . Since  $s^*$  is a non-increasing function of  $B$  (Proposition 5.3), the value of  $B$  which yields  $s^* = \min \left( M + 1, \left\lceil \frac{Np}{T} - \frac{1-p}{p} \right\rceil \right)$  can be found using a binary search algorithm.

□

## B Additional results when $P_{3G} \rightarrow \infty$

### B.1 Matrix form and linear program

$$A_1 = \begin{pmatrix} 1 & -1 & & & & \\ 0 & 1 & -1 & & & \\ 0 & 0 & 1 & -1 & & \\ & & & \ddots & & \\ & & & & 1 & -1 \\ & & & & & \ddots \\ & & & & & & 1 & -1 \\ & & & & & & 0 & 0 \end{pmatrix} \quad (104)$$

$$A_2 = \begin{pmatrix} -p+1 & -(1-p) & & & & \\ -p & 1 & -(1-p) & & & \\ -p & 0 & 1 & -(1-p) & & \\ & & & \ddots & & \\ & & & & 1 & -(1-p) \\ & & & & & \ddots \\ -p & & & & & & 1 & -(1-p) \\ -p & & & & & & 0 & -p \end{pmatrix} \quad (105)$$

Let  $V$  be the solution of the optimality equations.

$$A_1 V \geq U \quad (106)$$

$$A_2 V \geq U - G + pP - pB \quad (107)$$

Letting  $\widetilde{A}_1$  and  $\widetilde{A}_2$  denote matrices  $A_1$  and  $A_2$  with their last line and column removed, yields the following LP

$$\min \quad |\widetilde{V}| \tag{108}$$

$$\widetilde{V} \geq \widetilde{A}_1^{-1} U \tag{109}$$

$$\widetilde{V} \geq A_2^{-1}(U - G + pP - pB) \tag{110}$$

where  $\widetilde{V}$  is the vector  $V$  with its last element removed.

## C Derivation of Main Results Accounting for 3G and WiFi

We now analyze the case in which 3G is available. The optimality conditions are, for  $1 \leq x \leq M$ ,

$$\boxed{V(x) + g_u = \max \left( U(x) + K(x), U(x) - G + pV(1) + (1-p)K(x) - pP + pB, U(x) - G + V(1) - pP - (1-p)P_{3G} + B \right)} \tag{111}$$

where  $K(x) = V(\min(M, x+1))$ .

Let  $F(a, x)$  be the relative reward at state  $x$  when action  $a$  is chosen, plus  $g_u$ ,

$$F(0, x) = U(x) + V(\min(M, x+1)) \tag{112}$$

$$F(1, x) = U(x) - G + pV(1) + (1-p)V(\min(M, x+1)) - pP + pB \tag{113}$$

$$F(2, x) = U(x) - G + V(1) - pP - (1-p)P_{3G} + B. \tag{114}$$

The optimality conditions are, for  $x = 1, \dots, M$ ,

$$V(x) = \max (F(0, x) - g_u, F(1, x) - g_u, F(2, x) - g_u). \tag{115}$$

Let  $\mathcal{F}(a, x)$ ,  $a = 0, 1, 2$ , and  $x = 1, \dots, M$ , be boolean variables such that

$$\mathcal{F}(0, x) = \top \Leftrightarrow F(0, x) \geq \max_i (F(i, x)) \tag{116}$$

$$\mathcal{F}(1, x) = \top \Leftrightarrow F(1, x) \geq \max_i (F(i, x)) \tag{117}$$

$$\mathcal{F}(2, x) = \top \Leftrightarrow F(2, x) \geq \max_i (F(i, x)) \tag{118}$$

$$\tag{119}$$

From (112)-(114), a policy that selects  $a(x)$ ,  $x = 1, \dots, M$ , as follows, is optimal



$$\begin{cases} 0, & \text{if } V(1) - V(\min(M, x+1)) \leq G/p + P - B & \text{and } V(1) - V(\min(M, x+1)) \leq G + pP + (1-p)P_{3G} - B & (i) \\ 1, & \text{if } V(1) - V(\min(M, x+1)) \geq G/p + P - B & \text{and } V(1) - V(\min(M, x+1)) \leq P_{3G} - B & (ii) \\ 2, & \text{if } V(1) - V(\min(M, x+1)) \geq P_{3G} - B & \text{and } V(1) - V(\min(M, x+1)) \geq G + pP + (1-p)P_{3G} - B & (iii) \end{cases} \quad (120)$$

where ties are broken arbitrarily.

## C.1 Proposition C.1

**Proposition C.1.**  $V(x)$  is a non-increasing function.

*Proof.* According to (120)-(ii), if  $P_{3G} < G/p + P$ , there exists an optimal policy in which action 1 is not selected. Therefore, in what follows we separately consider two scenarios,  $P_{3G} \leq G/p + P$  and  $P_{3G} > G/p + P$ .

$$\boxed{\boxed{\text{scenario 1) } P_{3G} \leq \frac{G}{p} + P}} \quad (121)$$

In this scenario, there exists an optimal policy wherein  $a(x) \neq 1$ ,  $x = 1, \dots, M$ . We show by induction that

- $V(M-i-1) - V(M-i) \geq 0$  for  $i = 0, \dots, M-1$ ;
- $\mathcal{F}(M-i, 0) = \top \Rightarrow \mathcal{F}(M-i-1, 0) = \top$  for  $i = 0, \dots, M-1$ .

(120)(i) and (120)(iii) yield the following remark, used in the analysis of the base cases of the inductive arguments that follow.

**Remark C.1.**  $\mathcal{F}(M, 0) = \top \iff \mathcal{F}(M-1, 0) = \top$ . ■

We consider two cases,  $\mathcal{F}(M, 0) = \top$  and  $\mathcal{F}(M, 2) = \top$ .

- $\mathcal{F}(M, 0) = \top$ :

**Base case:** We first show that  $V(M-1) \geq V(M)$ . It follows from (115) and remark C.1 that

$$V(M-1) = U(M-1) + V(M) - g_u \quad (122)$$

$$\begin{aligned} V(M) &= U(M) + V(M) - g_u \\ &\geq U(M) + V(1) - pP - G - (1-p)P_{3G} + B - g_u. \end{aligned} \quad (123)$$

Hence, (122)-(123) together with the fact that  $U(M)$  is non-increasing and  $U(M) = 0$  yield

$$V(M-1) \geq V(M) \geq V(1) - pP - G - (1-p)P_{3G} + B - g_u.$$

**Induction hypothesis:** Assume that  $V(M-m-1) - V(M-m) \geq 0$  for  $m < t$  and that  $\mathcal{F}(M-m, 0) = \top$ , for  $m \leq t$ .

**Induction step:** Next, we show that  $V(M-t-1) - V(M-t) \geq 0$  and that  $\mathcal{F}(M-t-1, 0) = \top$ .

It follows from the induction hypothesis that

$$V(1) - pP - G + (1-p)P_{3G} + B \leq V(M-t+1) \leq V(M-t) \quad (124)$$

and  $\mathcal{F}(M-t, 0) = \top$ . Therefore,

$$V(1) - V(M-t) \leq V(1) - V(M-t+1) \leq pP + G + (1-p)P_{3G} - B \quad (125)$$

The rightmost inequality in (125) and optimality conditions (115) imply that  $\mathcal{F}(M-t-1, 0) = \top$ , which yield

$$V(M-t-1) = U(M-t-1) + V(M-t) - g_u \quad (126)$$

$$V(M-t) = U(M-t) + V(M-t+1) - g_u \quad (127)$$

Therefore,  $V(M-t+1) \leq V(M-t)$  (induction hypothesis) and  $U(M-t) \geq U(M-t-1)$  together with (126)-(127) yield  $V(M-t) \leq V(M-t-1)$ .

- $\mathcal{F}(M, 2) = \top$ :

**Base case:** We first show that  $V(M-1) \geq V(M)$ . It follows from the optimality conditions (115) and remark C.1 that

$$V(M-1) = U(M-1) + V(1) - G - pP - (1-p)P_{3G} - g_u + B \quad (128)$$

$$V(M) = U(M) + V(1) - G - pP - (1-p)P_{3G} - g_u + B. \quad (129)$$

Hence, (128)-(129) and  $U(M) < U(M-1)$  yield  $V(M-1) \geq V(M)$ .

**Induction hypothesis:** Assume that  $V(M-m-1) - V(M-m) \geq 0$  for  $m < t$  and that  $\mathcal{F}(M-m+1, 0) = \top \Rightarrow \mathcal{F}(M-m, 0) = \top$ , for  $m \leq t$ .

**Induction step:** Next, we show that  $V(M-t-1) - V(M-t) \geq 0$  and  $\mathcal{F}(M-t, 0) = \top \Rightarrow \mathcal{F}(M-t-1, 0) = \top$ . We consider three cases,

- $\mathcal{F}(M-t, 0) = \top$  This case is similar to the corresponding one for  $\mathcal{F}(M, 0) = \top$ .
- $\mathcal{F}(M-t, 2) = \top$  and  $\mathcal{F}(M-t-1, 2) = \top$  Next, we show that  $V(M-t-1) - V(M-t) \geq 0$ .

$$V(M-t) = U(M-t) + V(1) - G - pP - (1-p)P_{3G} - g_u + B \quad (130)$$

$$V(M-t-1) = U(M-t-1) + V(1) - G - pP - (1-p)P_{3G} - g_u + B. \quad (131)$$

Since  $U(M-t) \leq U(M-t-1)$ , it follows from (130)- (131) that  $V(M-t) \leq V(M-t-1)$ .

- $\mathcal{F}(M-t, 2) = \top$  and  $\mathcal{F}(M-t-1, 0) = \top$  Next, we show that  $V(M-t-1) - V(M-t) \geq 0$ .

$$V(M-t) = U(M-t) + V(1) - G - pP - (1-p)P_{3G} - g_u + B \quad (132)$$

$$V(M-t-1) = U(M-t-1) + V(M-t) \quad (133)$$

Since  $U(M-t-1) \geq 0$ , it follows from (133) that  $V(M-t) \leq V(M-t-1)$ .

$$\boxed{\text{scenario 2) } P_{3G} > \frac{G}{p} + P} \quad (134)$$

We show by induction that, for  $i = 0, \dots, M-1$ ,

- $V(M-i-1) \geq V(M-i)$ ,
- $\mathcal{F}(M-i, 0) = \top \Rightarrow \mathcal{F}(M-i-1, 0) = \top$
- $\mathcal{F}(M-i, 1) = \top \Rightarrow (\mathcal{F}(M-i-1, 1) = \top \text{ or } \mathcal{F}(M-i-1, 0) = \top)$

We consider three cases,  $\mathcal{F}(M, 0) = \top$ ,  $\mathcal{F}(M, 1) = \top$  and  $\mathcal{F}(M, 2) = \top$ .

(120)(i), (120)(ii) and (120)(iii), together, yield the following remarks, used in the analysis of the base cases of the inductive arguments that follow.

**Remark C.2.**  $\mathcal{F}(M, 1) = \top \iff \mathcal{F}(M-1, 1) = \top$ . ■

**Remark C.3.**  $\mathcal{F}(M, 2) = \top \iff \mathcal{F}(M-1, 2) = \top$ . ■

- $\mathcal{F}(M, 0) = \top$ : The proof is similar to that of scenario 1.
- $\mathcal{F}(M, 1) = \top$ :

**Base case:** It follows from remark C.2 and (112)-(115) that

$$V(M-1) = U(M-1) + pV(1) - pP + pB - (1-p)V(M) - g_u \quad (135)$$

$$V(M) = U(M) + pV(1) - pP + pB - (1-p)V(M) - g_u. \quad (136)$$

Hence, inequalities in (135)-(136) yield

$$V(M-1) \geq V(M)$$

**Induction hypothesis:** Assume that, for  $m = 0, \dots, t-1$ ,

- a)  $V(M-m-1) \geq V(M-m)$
- b)  $\mathcal{F}(M-m, 0) = \top \Rightarrow \mathcal{F}(M-m-1, 0) = \top$
- c)  $\mathcal{F}(M-m, 1) = \top \Rightarrow (\mathcal{F}(M-m-1, 1) = \top \text{ or } \mathcal{F}(M-m-1, 0) = \top)$

**Induction step:** Next, we show that

- a)  $V(M-t-1) \geq V(M-t)$
- b)  $\mathcal{F}(M-t, 0) = \top \Rightarrow \mathcal{F}(M-t-1, 0) = \top$
- c)  $\mathcal{F}(M-t, 1) = \top \Rightarrow (\mathcal{F}(M-t-1, 1) = \top \text{ or } \mathcal{F}(M-t-1, 0) = \top)$

It follows from the induction hypothesis that  $V(M-t+1) \leq V(M-t)$ . We consider three cases,

- $\mathcal{F}(M-t, 0) = \top$  The proof is similar to the corresponding one for  $H(M, 1) \leq H(M, 0)$  in the proof of Proposition A.1.
  - $\mathcal{F}(M-t, 1) = \top$  The proof is similar to the corresponding one for  $H(M, 1) \geq H(M, 0)$  in the proof of Proposition A.1.
  - $\mathcal{F}(M-t, 2) = \top$  Due to the induction hypothesis (cases b) and c)), one of the two cases above holds. Hence, it is not necessary to consider the case  $\mathcal{F}(M-t, 2) = \top$ .
- $\mathcal{F}(M, 2) = \top$ :

**Base case:** It follows from remark C.3 that  $V(M-1) \leq V(M)$ .

**Induction hypothesis:** Assume that, for  $m = 0, \dots, t-1$ ,

- $V(M-m-1) \geq V(M-m)$
- $\mathcal{F}(M-m, 0) = \top \Rightarrow \mathcal{F}(M-m-1, 0) = \top$
- $\mathcal{F}(M-m, 1) = \top \Rightarrow (\mathcal{F}(M-m-1, 1) = \top \text{ or } \mathcal{F}(M-m-1, 0) = \top)$

**Induction step:** Next, we show that

- $V(M-t-1) \geq V(M-t)$
- $\mathcal{F}(M-t, 0) = \top \Rightarrow \mathcal{F}(M-t-1, 0) = \top$
- $\mathcal{F}(M-t, 1) = \top \Rightarrow (\mathcal{F}(M-t-1, 1) = \top \text{ or } \mathcal{F}(M-t-1, 0) = \top)$

We consider the following cases.

- $\mathcal{F}(M-t, 0) = \top$  The proof is similar to that of scenario 1.
- $\mathcal{F}(M-t, 1) = \top$  First, we show that

$$(\mathcal{F}(M-t-1, 1) = \top) \text{ or } (\mathcal{F}(M-t-1, 0) = \top) \quad (137)$$

For the sake of contradiction, assume

$$(\mathcal{F}(M-t-1, 2) = \top) \text{ and } ((\mathcal{F}(M-t-1, 0) \neq \top) \text{ and } (\mathcal{F}(M-t-1, 1) \neq \top)) \quad (138)$$

Then,

$$F(M-t-1, 2) \geq F(M-t-1, 1) \quad (139)$$

$$F(M-t, 1) \geq F(M-t, 2) \quad (140)$$

Replacing (112)-(114) into (139)-(140),

$$\begin{aligned} -G + V(1) - pP - (1-p)P_{3G} + B &\geq -G + pV(1) + (1-p)V(M-t) - pP + (1-p)P_{3G} \\ -G + pV(1) + (1-p)V(M-t+1) - pP + pB &\geq -G + V(1) - pP - (1-p)P_{3G} + B \end{aligned} \quad (142)$$

From (141) and (142) we obtain, respectively,

$$V(1) - V(M-t) - P_{3G} + B \geq 0 \quad (143)$$

$$-V(1) + V(M-t+1) + P_{3G} - B \geq 0 \quad (144)$$

Applying the induction hypothesis to (143),

$$V(1) - V(M-t+1) - P_{3G} + B \geq 0 \quad (145)$$

Therefore, from (144) and (145),

$$V(M-t+1) + P_{3G} - B = V(1) \quad (146)$$

(146) together with (112)-(114) imply that both  $\mathcal{F}(M-t+1, 2) = \top$  and  $\mathcal{F}(M-t+1, 1) = \top$  are optimal, which contradicts (138).

To show that  $V(M-t-1) \geq V(M-t)$ , an argument similar to the corresponding one for  $H(M, 1) \geq H(M, 0)$  in the proof of Proposition A.1 can be applied.

- $\mathcal{F}(M-t, 2) = \top$  We consider three subcases.

- \*  $\mathcal{F}(M-t, 2) = \top$  and  $\mathcal{F}(M-t-1, 2) = \top$  The proof is similar to the corresponding case in

scenario 1.

\*  $\mathcal{F}(M-t, 2) = \top$  and  $\mathcal{F}(M-t-1, 0) = \top$  The proof is similar to the corresponding case in scenario 1.

\*  $\mathcal{F}(M-t, 2) = \top$  and  $\mathcal{F}(M-t-1, 1) = \top$  In this case, it follows from  $F(1, M-t-1) \geq F(2, M-t-1)$  that

$$-G + pV(1) + (1-p)V(M-t) - pP + pB \geq -G + V(1) - pP - (1-p)P_{3G} + B \quad (147)$$

From (147),

$$pV(1) + (1-p)V(M-t) + pB \geq V(1) - (1-p)P_{3G} + B \quad (148)$$

Therefore,

$$U(M-t-1) + pV(1) + (1-p)V(M-t) + pB \geq U(M-t) + V(1) - (1-p)P_{3G} + B \quad (149)$$

(149) together with (112)-(114) imply that  $V(M-t-1) \geq V(M-t)$ .

□

## C.2 Proposition 3.1

*Proof.* The proof follows from Proposition C.1 and optimality conditions (120). □

## C.3 Proposition C.2 (conditions for the optimal policy to consist of using only one action)

**Proposition C.2.** *Let  $a(x) = \alpha$ ,  $x = 1, \dots, M$  be an optimal policy.*

*If  $P_{3G} \leq \frac{G}{p} + P$  and*

- $\alpha = 2$ ,

$$U(1) - U(2) \geq G + (1-p)P_{3G} + pP - B. \quad (150)$$

- $\alpha = 0$ ,

$$\sum_{j=1}^{M-1} U(M-j) \leq G + (1-p)P_{3G} + pP - B. \quad (151)$$

*If  $P_{3G} > \frac{G}{p} + P$  and*

- $\alpha = 0$ ,

$$\sum_{j=1}^{M-1} U(M-j) \leq \frac{G}{p} + P - B. \quad (152)$$

- $\alpha = 1$ ,

$$U(1) - p \sum_{i=1}^{M-1} U(i)(1-p)^{i-1} \geq (1-p)(G/p + P - B) \quad (153)$$

$$\sum_{i=1}^{M-1} U(i)(1-p)^{i-1} < P_{3G} - B. \quad (154)$$

- $\alpha = 2$ ,

$$U(1) - U(2) \geq P_{3G} - B. \quad (155)$$

*Proof.* The proof is similar in spirit to the one of Proposition 5.1.  $\square$

**Remark C.4.** *If the optimal threshold policy is unique, the conditions in Proposition C.2 are necessary and sufficient for the optimal policy to satisfy (150)- (155).*

## C.4 Optimal threshold properties with 3G

To derive the optimal threshold strategy properties, we distinguish two scenarios:  $P_{3G} \leq \frac{G}{p} + P$  and  $P_{3G} > \frac{G}{p} + P$ .

$$\boxed{\boxed{\text{scenario 1) } P_{3G} \leq \frac{G}{p} + P}} \quad (156)$$

In this scenario,

$$a(x) = \begin{cases} 0, & \text{if } x < s_{3G} \\ 2, & \text{otherwise} \end{cases} \quad (157)$$

The transition matrix of the system,  $\mathcal{P}_u$ , is

$$\mathcal{P}_u = \begin{pmatrix} 0 & 1 & & & & & \\ 0 & 0 & 1 & & & & \\ 0 & 0 & 0 & 1 & & & \\ & & & \ddots & & & \\ 1 & & & & 0 & 0 & \\ & & & & & \ddots & \\ 1 & & & & & & 0 & 0 \\ 1 & & & & & & & 0 \end{pmatrix} \quad (158)$$

Let  $\pi$  be the steady state solution of the system for a threshold  $s_{3G}$ ,  $\pi \mathcal{P}_u = \pi$ . Then,

$$\pi_1 = \pi_i = \frac{1}{s_{3G}}, \quad i = 1, \dots, s_{3G} \quad (159)$$

$$\pi_i = 0, \quad i = s_{3G} + 1, \dots, M \quad (160)$$

$$(161)$$

The expected age,  $A(s_{3G}) = \sum_{i=1}^M \pi_i i$ , is

$$A(s_{3G}) = \begin{cases} (s_{3G} + 1)/2, & \text{if } 1 \leq s_{3G} \leq M \\ M, & \text{if } s_{3G} = M + 1 \end{cases} \quad (162)$$

The average reward of this threshold policy is

$$E[r; s_{3G}] = \begin{cases} \frac{1}{s_{3G}} [\sum_{i=1}^{s_{3G}} U(i) - G - pP - (1-p)P_{3G} + B], & \text{if } 1 \leq s_{3G} \leq M \\ 0, & \text{if } s_{3G} = M + 1. \end{cases} \quad (163)$$

The optimal threshold value  $s_{3G}^*$  is

$$s_{3G}^* = \min \left\{ s \mid \sum_{i=1}^s d(i) - sU(s+1) \geq G + pP + (1-p)P_{3G} \right\} \quad (164)$$

$$\boxed{\boxed{\text{scenario 2) } P_{3G} > \frac{G}{p} + P}} \quad (165)$$

In this scenario, let  $s_W = s$ ,

$$a(x) = \begin{cases} 0 & \text{if } x < s_W \\ 1 & \text{if } s_W \leq x < s_{3G} \\ 2 & \text{otherwise} \end{cases} \quad (166)$$



The case  $s_{3G} = M + 1$  was considered in Appendix A. Next, we consider the case  $s_{3G} \leq M$ . The transition matrix of the system,  $\mathcal{P}_u$ , is

$$\mathcal{P}_\pi = \begin{pmatrix} 0 & 1 & & & & & & & \\ 0 & 0 & 1 & & & & & & \\ 0 & 0 & 0 & 1 & & & & & \\ & & & \ddots & & & & & \\ p & & & & 0 & 1-p & & & \\ & & & & \ddots & & & & \\ p & & & & & & 0 & 1-p & \\ 1 & & & & & & & 0 & \\ 1 & & & & & & & 0 & \\ 1 & & & & & & & 0 & \end{pmatrix} \quad (167)$$

The steady state solution,  $\pi$ , of the above system is

$$\pi_1 = \pi_i, \quad i = 1, \dots, s_W \quad (168)$$

$$\pi_i = \pi_{i-1}(1-p) = \pi_1(1-p)^{i-s_W}, \quad i = s_W + 1, \dots, s_{3G} \quad (169)$$

$$\pi_i = 0, \quad i = s_{3G} + 1, \dots, M \quad (170)$$

$$\sum_{i=1}^M \pi_i = 1 \quad (171)$$

To obtain  $\pi_1$ , we sum (168) and (169),

$$s_W \pi_1 + \pi_1 \sum_{i=s_W+1}^{s_{3G}} (1-p)^{i-s} = 1 \quad (172)$$

therefore,

$$\pi_1 = \frac{1}{s_W - 1 + (1 - (1-p)^{s_{3G}-s_W+1})p^{-1}}. \quad (173)$$

The expected reward is

$$E[r; s_W, s_{3G}] = \pi_1 \left[ \sum_{i=1}^{s_W-1} U(i) + \sum_{i=s_W}^{s_{3G}} (1-p)^{i-s_W} U(i) - (G/p + P - B) - (P_{3G} - P - G/p)(1-p)^{s_{3G}-s_W+1} \right]. \quad (174)$$

The optimal thresholds  $(s_W^*, s_{3G}^*)$  are  $(s_W^*, s_{3G}^*) = \arg \max_{(s_W, s_{3G})} \{E[r; s_W, s_{3G}]\}$ .

## C.5 On the number of optimal thresholds

If there exists an optimal policy which consists of switching being inactive and WiFi, the results presented in §A.4 hold if we restrict to policies for which  $a(x) \in \{0, 1\}$  ( $x = 1, \dots, M$ ). In contrast, if there is an optimal policy which consists of adopting 3G at state  $s_{3G} < M$ , the optimal policy is in general not unique, since changing the actions chosen at states  $s > s_{3G}$  does not compromise the optimality of the resulting policy. A characterization of the number of optimal two-threshold policies in this context is subject of future work.

## D Learning algorithm

### D.1 Proof of Proposition 6.2

In what follows, we assume that the PUBLISHER PROBLEM admits a solution  $B^*$ , and that  $B^* > 0$ .

*Proof.* We index the discrete rounds of Algorithm 1 by  $n$  ( $n = 1, 2, 3, \dots$ ),

$$B_{n+1} \leftarrow \min \left( \widehat{B}, \max \left( 0, B_n + \frac{\alpha}{n} (T - Q_n) \right) \right) \quad (175)$$

Then, we prove a generalized version of Proposition 6.2. Namely, we consider the following algorithm, obtained from (175) after removing the max and min operators,

$$B_{n+1} \leftarrow B_n + \frac{\alpha}{n} (T - Q_n) \quad (176)$$

Next, we show that (176) converges to the optimal solution of the PUBLISHER PROBLEM with probability 1. If  $\widehat{B}$  is sufficiently large, the convergence of (176) to the optimal solution of the PUBLISHER PROBLEM implies the convergence of (175) to such a solution as well (see [40, Lemma 3.3.8 and §5.4] for details).

#### D.1.1 Martingale definition

Let  $M_{n+1} = E[Q_n] - Q_n$  and  $g(B) = T - E[Q_n(B)]$ . Then

$$B_{n+1} \leftarrow B_n + \frac{\alpha}{n} (g(B_n) + M_{n+1}) \quad (177)$$

The history of values  $\{B_m\}$  and  $\{M_m\}$ ,  $m = 1, \dots, n$ , yields a  $\sigma$ -field  $\mathcal{F}_n$ ,  $\mathcal{F}_n = \sigma(B_m, M_m, m \leq n)$ .

Note that  $\{B_m\}$ ,  $m = 2, \dots, n$  are fully characterized by  $B_1$  and  $\{M_m\}$ ,  $m = 1, \dots, n$ . Therefore  $\mathcal{F}_n = \sigma(B_1, M_m, m \leq n)$ .

Moreover,  $\{M_n\}$ ,  $n = 1, 2, \dots$  is a sequence of zero mean random variables satisfying

$$E[M_{n+1} | \mathcal{F}_n] = 0 \text{ a.s., } n \geq 0.$$

As a consequence,  $\{M_n\}$  is a martingale difference sequence with respect to the increasing  $\sigma$ -fields  $\mathcal{F}_n$ ,  $n = 1, 2, \dots$

Recall that  $N$  is the number of mobiles in the network. Note that

$$Q_n \leq N \quad (178)$$

Therefore,  $|M_{n+1}| = |E[Q_n] - Q_n| \leq N$ , thus  $\{M_n\}$  is square-integrable, i.e.,

$$E[|M_{n+1}|^2 | \mathcal{F}_n] \text{ is finite} \quad (179)$$

### D.1.2 Differential inclusion definition

Since the function  $g(B) = T - E[Q_n(B)]$  is discontinuous, we use differential inclusions, an extension of the ordinary differential equations, to approximate (176). The solution concept adopted for the differential inclusions is the *Filippov solution* [41].

Let  $s_0$  be the number of discontinuity points of  $g$  in the interval  $(0, \widehat{B}]$ . Let  $B^{(i)}$  be the value of  $B$  at the  $i^{th}$  discontinuity point of  $g(B)$ ,  $i = 0, 1, 2, \dots, s_0$ . Then,  $B^{(0)} = 0 < B^{(1)} < \dots < B^{(s_0)} = \widehat{B}$  (see Figure 8).

Next, we approximate (176) by the following differential inclusion

$$\dot{B}(t) \in F(B(t)) \quad (180)$$

where  $F$  is a set-valued map

$$F : [0, \widehat{B}] \rightarrow \text{subset of } \mathbb{R}$$

$F$  is defined as follows

$$F(B) = \begin{cases} \{g(B)\}, & \text{if } B \neq B^{(i)}, \quad i = 1, \dots, s_0 \\ [g(B^{(i-1)}), g(B^{(i)})], & \text{if } B = B^{(i)}, \quad i = 1, \dots, s_0 \end{cases} \quad (181)$$

Since the function  $F$  is semi-continuous and, for all  $B(t)$ ,  $F(B(t))$  is a compact and a convex set, there exists a solution to the differential inclusion (180). In addition, it follows from Theorem 2 in [41] that such a solution is unique.

Let  $B^*$  be the following solution to the PUBLISHER PROBLEM,

$$B^* = \max \left\{ B^{(i)} \mid g(B^{(i)}) \geq 0 \right\} \quad (182)$$

Note that  $0 \in [g(B^{(i)}) = g(B^*), g(B^{(i+1)})]$  and let

$$\tilde{T} = \max \left\{ t \leq T \mid t - \frac{N}{s(B^*) + (1-p)/p} = 0 \right\} \quad (183)$$

The *basin of attraction* of (180) is (see Figure 8),

$$\begin{cases} [B^{(i)} = B^*, B^{(i+1)}], & \text{if } T = \tilde{T} \\ B^{(i+1)}, & \text{if } T \neq \tilde{T}, \text{ where } B^{(i)} = B^* \end{cases}$$

We establish the global stability of the basin of attraction of (180) in §D.1.3, and then in §D.1.4 we use [40, Theorem 2, §5] to show that the algorithm (176) converges to the basin of attraction of (180).

### D.1.3 Global stability of basin of attraction

In order to establish the global stability of the basin of attraction of (180), we use the following Lyapunov function

$$V(B) = (B - B^*)^2 \quad (184)$$

We begin by stating two definitions which generalize the notions of gradients and derivatives to encompass differential inclusions [42].

**Definition D.1.** *The Clarke generalized gradient of  $V$  in  $B$  is defined as*

$$\partial V(B) = \text{co}\left\{\lim_{l \rightarrow \infty} \Delta V(B_l) : (B_l) \rightarrow (B)\right\} \quad (185)$$

where  $\text{co}\{A\}$  is the smallest convex and bounded set containing  $A$ .

**Definition D.2.** *The set-valued derivative of  $V$  is defined as*

$$\dot{V}(B) = \{a \in \mathbb{R} : \exists v \in F(B) \text{ so that } pv = a, \forall p \in \partial V(B)\}$$

It follows from the two definitions above that

$$\partial V(B) = \{2(B - B^*)\} \quad (186)$$

and

$$\dot{V}(B) = \begin{cases} \{2(B - B^*)g(B)\}, & \text{if } B \neq B^{(j)}, \quad j = 1, \dots, s_0 \\ [2(B - B^*)g(B^{(j-1)}), 2(B - B^*)g(B^{(j)})], & \text{if } B = B^{(j)} \end{cases}$$

We divide the analysis of the stability of the basin of attraction of (180) into two cases, varying according to whether  $T \neq \tilde{T}$  (see Figure 8(b)) or  $T = \tilde{T}$  (see Figure 8(c))

- If  $T \neq \tilde{T}$  (see Figure 8(b)),  $\dot{V}(B) < 0$  for  $B \neq B^*$ ,
  - $B \neq B^{(j)}$ : If  $B > B^*$ , then  $g(B) = T - E[Q_n(B)] < 0$  since  $E[Q_n]$  is an increasing function of  $B$  (see Proposition 5.3). If  $B < B^*$  the result follows similarly;
  - $B = B^{(j)}$  and  $B \neq B^*$ : If  $B > B^*$ , then  $g(B^{(j-1)}) = T - E[Q_n(B^{(j-1)})] < 0$  since  $E[Q_n]$  is an increasing function of  $B$  (see Proposition 5.3) and  $B^{(j-1)} \leq B^{(j)}$ . If  $B < B^*$  the result follows similarly.
- Using similar arguments, we can establish the stability of the basin of attraction when  $T = \tilde{T}$  (see Figure 8(c)). Let  $\underline{B}^* = B^* = B^{(i)}$  and  $\overline{B}^* = B^{(i+1)}$ . Then,  $\dot{V}(B) < 0$  for  $B \notin [g(\underline{B}^*), g(\overline{B}^*)]$ .

Therefore, the global stability of the basin of attraction of (180) follows from the Lyapunov theorem (see [40]).

#### D.1.4 Algorithm convergence to differential inclusion

Next, our goal is to show that (176) is well approximated by (180). To this goal, we define the linear interpolation of  $B_n$  and establish the finiteness of  $\sup_n |B_n|$ .

*Interpolation of  $\{B_n\}$ :* Let  $\bar{B}(t)$  be the linear interpolation of  $B_n$ ,  $n = 1, 2, 3, \dots$

$$\bar{B}(t) = \begin{cases} B_n, & t = \sum_{i=1}^n \alpha/i \\ B_n + \frac{t - \sum_{i=1}^n \alpha/i}{\sum_{i=1}^{n+1} \alpha/i - \sum_{i=1}^n \alpha/i} (B_{n+1} - B_n), & t \in (\sum_{i=1}^n \alpha/i, \sum_{i=1}^{n+1} \alpha/i) \end{cases} \quad (187)$$

$\bar{B}(t)$  is a piecewise linear and continuous function.

*Finiteness of  $\sup_n |B_n|$ :* Next, we show that  $\sup_n |B_n| < \infty$ . For  $B \in [0, \infty)$ ,  $F(B)$  is a compact and convex set and

$$\sup_{y \in F(B)} |y| < g(B^{(0)}) < g(B^{(0)})(1 + B) \quad (188)$$

Let  $h_c(B)$  be defined as in [40, §3.2-(A5)],  $h_c(B) = g(cB)/c = (T - N/(s(cB) + (1 - p)/p))/c$ . Then, for  $B \in [0, \infty)$ ,

$$h_\infty(B) = 0 \quad (189)$$

Therefore, similar arguments as those in [40, §3.2] applied to (179), (188) and (189) yield

$$\sup_n |B_n| < \infty \quad (190)$$

*Concluding the proof:* Given that (190) (which corresponds to [40, (5.2.2)]) holds, Theorem 2 in [40, §5] establishes that almost surely every limit point of (187) satisfies (180) (see [40] for details). As a consequence, since the sequence  $\{B_n\}_{n=1,2,\dots}$  generated by (176) is contained in (187), it converges almost surely to (180) (see Corollary 4 of [40, §5] for details).  $\square$

## D.2 Binary search and stochastic learning

In this paper we proposed the use of binary search to obtain the optimal bonus level when the publishers have complete information (see Proposition 6.1) and stochastic learning in face of incomplete information (see Proposition 6.2). A variation of binary search, as proposed in [43], can also be used in face of incomplete information. However, we do not have a proof of the convergence of the modified binary search algorithm in our setting. Since the binary search algorithm may have faster convergence to the optimal solution, it can be used to obtain an initial value for the stochastic learning, whose convergence was established

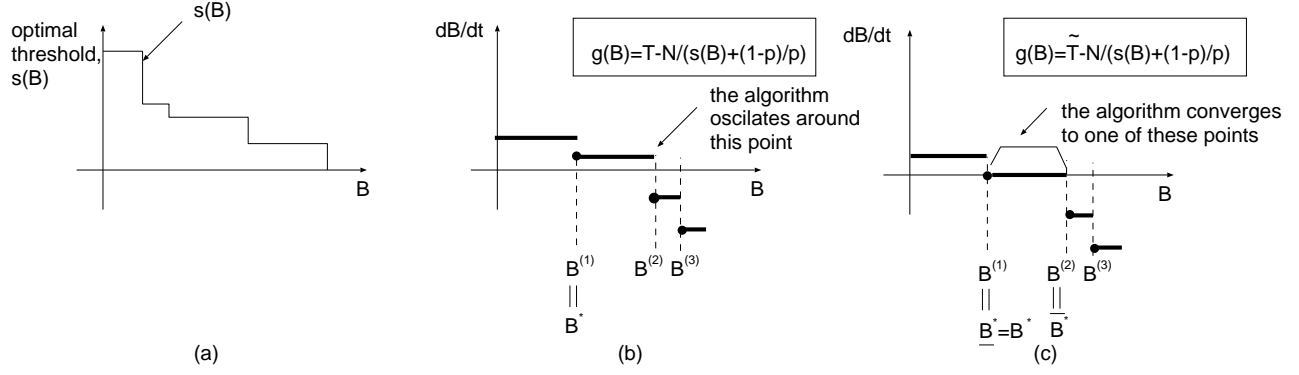


Figure 8: Convergence of learning algorithm. (a) the threshold policy adopted by users is piecewise constant; (b)  $g(B)$  when  $T \neq \tilde{T}$ ; (c)  $g(B)$  when  $T = \tilde{T}$

in Proposition 6.2. A comparison of pros and cons of the modified binary search and stochastic learning is provided in [43].

### D.3 Simulation

In order to 1) study the behavior of our algorithm when the population size varies and to 2) investigate how the convergence speed may be affected by the correlation among users, we conducted trace-driven simulations, whose results are shown in <http://www-net.cs.umass.edu/~sadoc/agecontrol/bus-ap-ct20/index.html>.

**Simulation setup** We set the parameters of our learning algorithm as follows:  $M=30$ ,  $G=0.4$ ,  $T=11$ ,  $P=100$  and  $\tau = 10$  time slots. The number of users is initially 105 and drops to 90 at round 100. Finally,  $\alpha=10$  for trace driven simulations, and  $\alpha = 20$  under the uniformity and independence assumption.

In our trace drive simulations in this section, we consider half of the population in one bus and the other half in another. Recall that for each bus shift we generate a string of zeros and ones corresponding to slots with and without a useful contact opportunity, respectively. In order to distinguish the users in two buses, while simulating the opportunities observed by each user we assume that half of them take their first observations in the beginning of the string of zeros and ones while the other half take their first observations in the middle of the trace. After taking their first observations, the subsequent ones are drawn in sequence out of the string of zeros and ones (our simulator is available at <http://www-net.cs.umass.edu/~sadoc/agecontrol/learning.tgz>).

Despite the correlations among users, for the 88 bus shifts analyzed our trace driven results did not significantly deviate from the ones obtained with uncorrelated users. In <http://www-net.cs.umass.edu/~sadoc/agecontrol/bus-> we show for each bus shift the results obtained with trace driven simulations as well as under the uniformity and independence assumption (the contact opportunity being obtained from traces). In what follows, we illustrate our results with two examples.

**Simulation Results** Figures 9 and 11 illustrate our simulation results for two bus shifts. Figures 9(a) and 11(a) show the contact opportunities. Note that while in Figure 9(a) the opportunities are roughly

uniform across different time slots, in Figure 11(a) there is a high concentration of opportunities close to Haigis Mall, but not many in-between arrivals at Haigis Mall. This, in turn, impacts the results obtained with the learning algorithms, as shown in Figures 9(b)-(c) and 11(b)-(c). The dotted lines represent the optimal range of bonuses. While in Figure 9(b) the bonus computed using the learning algorithm converges to optimal values, in Figure 11(b) the difference between the optimal bonus and the one obtained with the learning algorithm was around 20. Despite such difference, though, note that the number of transmissions experienced by the service providers oscillated around its target, eleven, in both cases (see Figures 9(c) and Figures 11(c)).

In Figure 9(c) the number of transmissions remains stable during certain intervals of time (for instance, in the interval  $[20,30]$ ). This is partly due to the synchronization of the users, that are assumed to begin with the same initial state and to be experience correlated contact opportunities. In contrast, if we consider users that experience contact opportunities uniformly at random, such synchronizations do not occur and the bonus computed using the learning algorithm converges to optimal values in both Figure 9 and 11 (see Figures 9(d)-(e) and Figures 11(d)-(e)).

Further insight into the synchronization among users in Figures 9(c) is obtained from Figure 10. Each curve in Figure 10 shows the age of the users in each of the buses. Since all the users are assumed to begin with age 1 at slot 1, the age of all users in a bus remains equal throughout our simulations. Figure 10 shows that from slots 300 to 400 (which correspond to rounds 30 to 40), in each round each user issues one update. Since there are 105 users, the average number of updates per slot is 10.5. Note that such synchronization does not occur between slots 400 and 450, when the number of updates issued by a user in a given slot varies between 1 and 2. The synchronization between users is an artifact due to the assumption that they have the same initial state at slot 1, and does not occur if users have their initial states sampled uniformly at random.



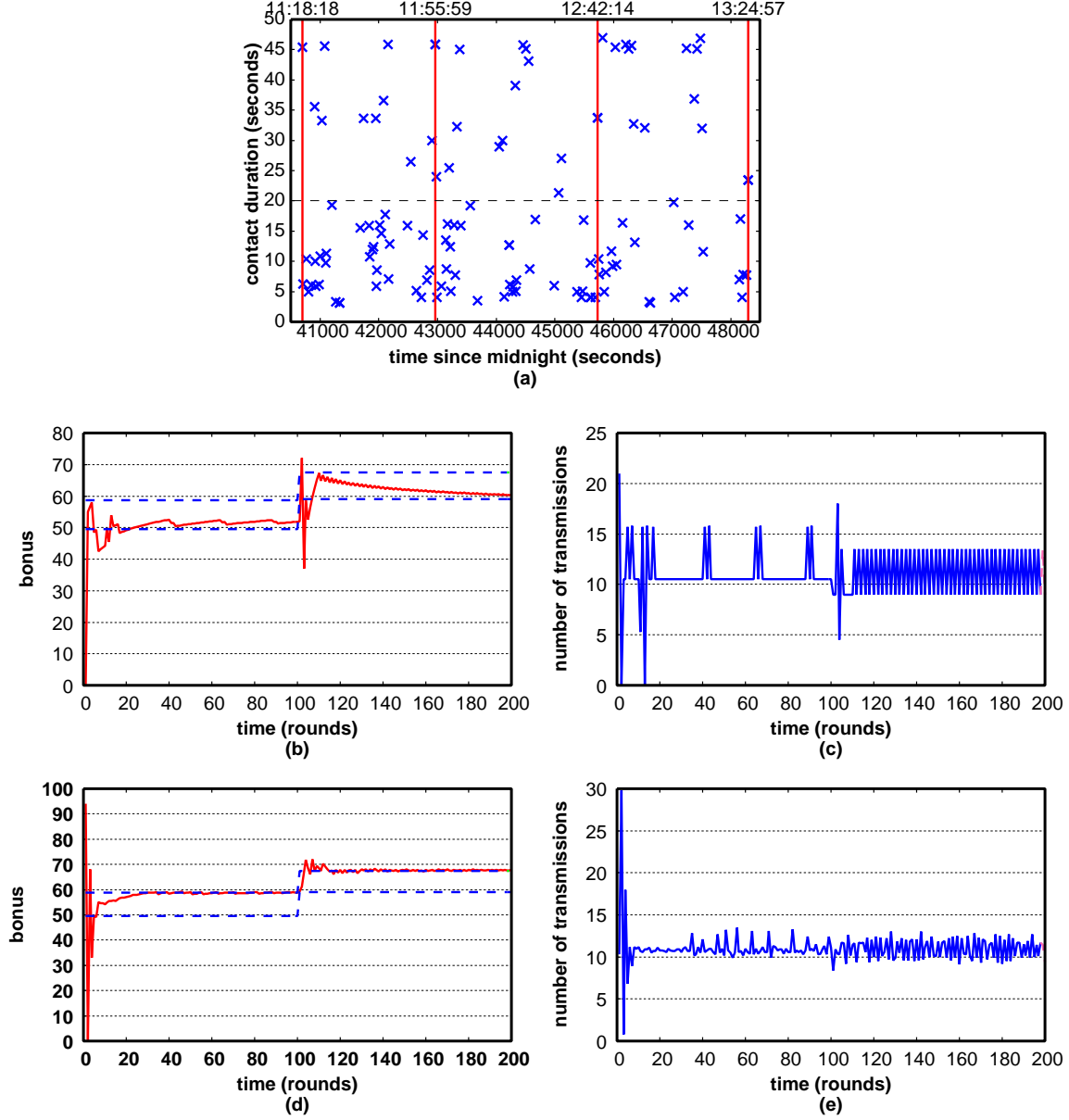


Figure 9: (a) Contact opportunities; (b) and (c) show the bonus and number of transmissions obtained from trace driven simulations; (d) and (e) show the bonus and number of transmissions obtained from simulations under uniformity and independence assumptions.

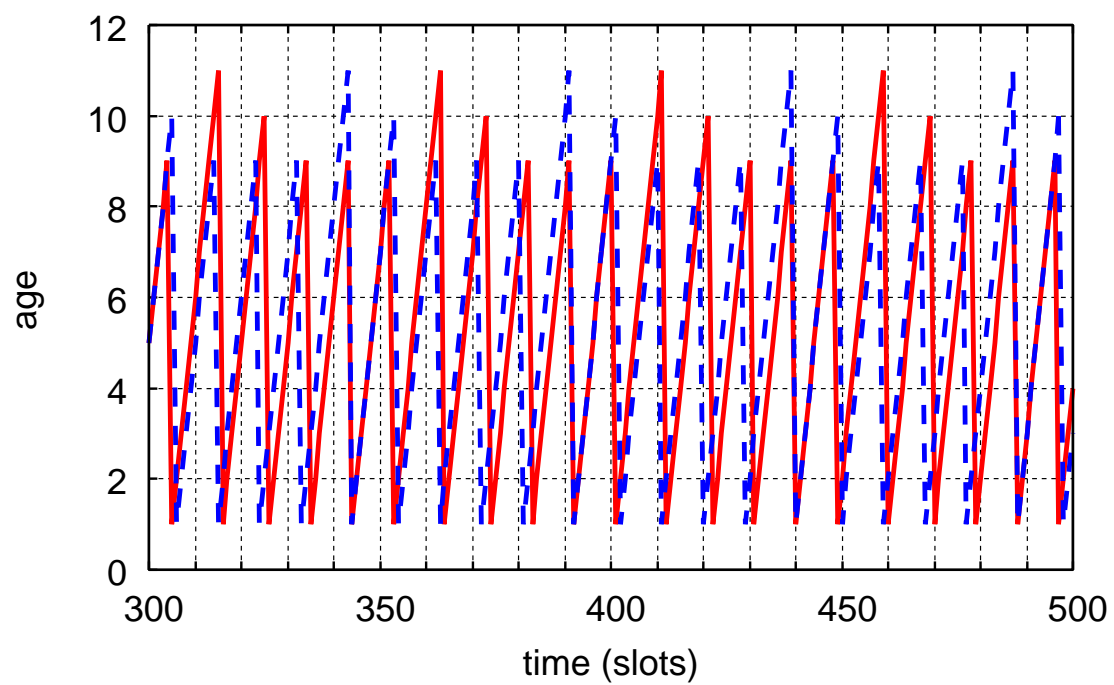


Figure 10: Sample path of age as a function of slot (trace driven simulation) See also Figure 9(c).

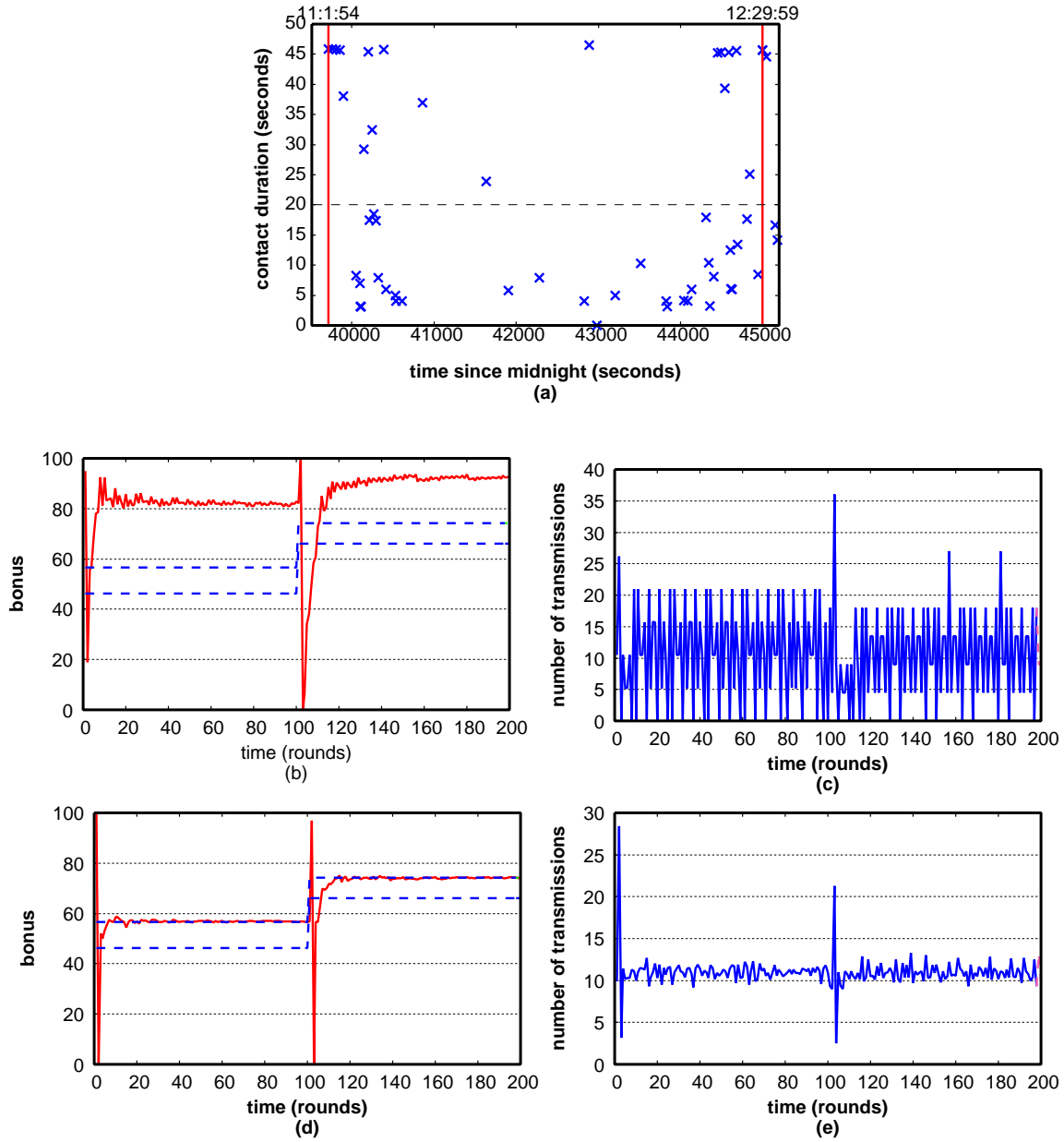


Figure 11: (a) Contact opportunities; (b) and (c) show the bonus and number of transmissions obtained from trace driven simulations; (d) and (e) show the bonus and number of transmissions obtained from simulations under uniformity and independence assumptions.

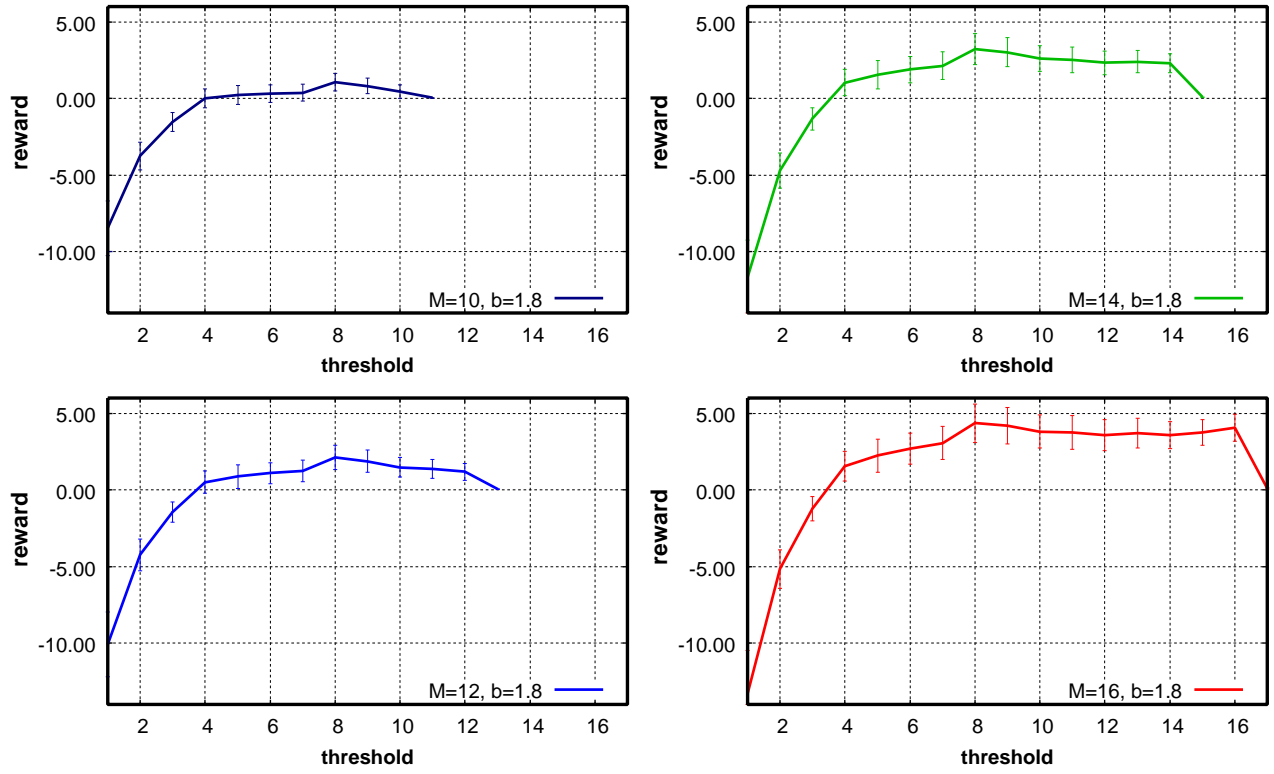


Figure 12: Trace-driven reward as a function of the threshold, where threshold is assumed the same at all bus shifts (Figure 4(c) with 95% confidence intervals)